

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 23, 2005, 23:11:27 ; Search time 230 Seconds
(without alignments)
1165.655 Million cell updates/sec

Title: US-10-763-042-9
Perfect score: 1962
Sequence: 1 MTFRSLQEMRSPHRVHA.....SAFQAFCLSSFFETRIACE 380

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1962	100.0	380	Q9SQUL_ARATH	Q9squl arabidopsis
2	1442.5	73.5	389	Q6NQP1_ARATH	Q6npq1 arabidopsis
3	1414.5	72.1	380	Q6YNN7_ARATH	Q6ynn7 arabidopsis
4	1149	58.6	406	Q8VY21_ARATH	Q8vy21 arabidopsis
5	1147.5	58.5	407	Q82257_ARATH	Q82257 arabidopsis
6	1139	58.1	406	Q8H485_ORYSA	Q8h485 oryza sativ
7	1015	51.7	451	Q69U54_ORYSA	Q69u54 oryza sativ
8	1013.5	51.7	445	Q75HX5_ORYSA	Q75hx5 oryza sativ
9	1013.5	51.7	455	Q9ZP59_ARATH	Q9zpz59 arabidopsis
10	999.5	50.9	445	Q9FRH7_ARATH	Q9frh7 arabidopsis
11	992	50.6	448	Q8LTA9_ORYSA	Q8lta9 oryza sativ
12	989	50.4	462	Q7XSU4_ORYSA	Q7xsu4 oryza sativ
13	989	50.4	462	Q9ST78_ORYSA	Q9st78 oryza sativ
14	969	49.4	440	Q53PP5_ORYSA	Q53pp5 oryza sativ
15	966.5	49.3	394	Q8GVE5_ARATH	Q8gve5 arabidopsis
16	957.5	48.8	394	Q7Y213_ARATH	Q7y213 arabidopsis
17	950.5	48.4	429	Q94483_ARATH	Q94483 arabidopsis
18	937	47.8	428	Q9SX13_ARAE	Q9sx13 lemna pauci
19	933.5	47.6	411	Q9M3H9_CICAR	Q9m3h9 cicor ariet
20	929.5	47.4	415	Q84JM8_ARATH	Q84jm8 arabidopsis
21	929	47.3	388	Q9C6B4_ARATH	Q9c6b4 arabidopsis
22	927.5	47.3	413	Q84U70_ARALY	Q84u70 arabidopsis
23	927.5	47.3	415	Q84UG3_ARATH	Q84ug3 arabidopsis
24	925.5	47.2	415	Q84UG2_ARATH	Q84ug2 arabidopsis
25	925.5	47.2	428	Q6ZZG9_ORYSA	Q6zzg9 oryza sativ
26	919.5	46.9	386	Q9ZFW1_ARATH	Q9zfw1 arabidopsis
27	843	43.0	379	Q93V18_ARATH	Q93v18 arabidopsis
28	832	42.4	356	Q5QM27_ORYSA	Q5qm27 oryza sativ
29	816.5	41.6	579	Q9MAG9_ARATH	Q9mag9 arabidopsis
30	810.5	41.3	372	Q68Y48_ORYSA	Q68y48 oryza sativ
31	800.5	40.8	352	Q688Y7_ORYSA	Q688y7 oryza sativ

RESULT 1
Q9SQUL_ARATH PRELIMINARY; PRT; 380 AA.
AC Q9SQUL;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE F24PI7.15 protein (Tubby-like protein TULP9).
GN Name=F24PI7.15; Synonyms=TULP9;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Lin X., Kaul S., Town C.D., Benito M.-I., Creasy T.H., Haas B.,
RA Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15064372; DOI=10.1104/pp.103.037820;
RA Lai C.P., Lee C.L., Chen P.H., Wu S.H., Yang C.C., Shaw J.P.;
RT "Molecular analyses of the Arabidopsis TUBBY-like protein gene family.";
RL Plant Physiol. 134:1586-1597(2004).
DR EMBL; AC011623; AAF08576.1; -; Genomic_DNA.
DR EMBL; AF487270; AAQ06243.1; -; mRNA.
DR HSSP; P50586; 117E.
DR GO; GO:0045449; P:regulation of transcription; TAS.
DR InterPro; IPR001810; F-box.
DR InterPro; IPR000007; Tubby.
DR Pfam; PF00646; F-box; 1.
DR Pfam; PF01167; Tub; 1.
DR PRINTS; PR01573; SUPERTUBBY.
DR PROSITE; PS01200; TUB 1; UNKNOWN 1.
SQ SEQUENCE 380 AA; 42310 MW; EDB57859F8B5F51D CRC64;

Query Match 100.0%; Score 1962; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 5,3e-165;
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MTFRSLQEMRSPHRVHA	AAASTANSSDPFSWSELPEELLREILIRVETDGGDWPSRR	60
Db	1	MTFRSLQEMRSPHRVHA	AAASTANSSDPFSWSELPEELLREILIRVETDGGDWPSRR	60
QY	61	NNVACAGVCRWRILTKEL	VAIVPEPSSKLTFFISLKSQSPRSLVQCFTKRNRNTQSYHL	120
Db	61	NNVACAGVCRWRILTKEL	VAIVPEPSSKLTFFISLKSQSPRSLVQCFTKRNRNTQSYHL	120
QY	121	YLGLTTLTDNGKFLAAS	KLKRAICTDYIIISLRSDDISKRSNAYLGRMRSNFLGTKFTV	180
Db	121	YLGLTTLTDNGKFLAAS	KLKRAICTDYIIISLRSDDISKRSNAYLGRMRSNFLGTKFTV	180

```
QY 181 FDGSOTGAAMOKRSNFIKVSPPVQSGPYPIAHISYELNVLGSRGPRMRCIMDTIPM 240
Db 181 FDGSOTGAAMOKRSNFIKVSPPVQSGPYPIAHISYELNVLGSRGPRMRCIMDTIPM 240
QY 241 SVESRGVVASTSISSFSRSSPVFRSHKPLRSNSASCDSGNNLGDPLVLNKAAPRW 300
Db 241 SVESRGVVASTSISSFSRSSPVFRSHKPLRSNSASCDSGNNLGDPLVLNKAAPRW 300
QY 301 HEQLRCWCLNFRGRVTVASVKQFQVAVSDCEAGQTSERIILQFGKVGKDMFTMDYGYPI 360
Db 301 HEQLRCWCLNFRGRVTVASVKQFQVAVSDCEAGQTSERIILQFGKVGKDMFTMDYGYPI 360
QY 361 SAFAQFAICLSSFETRIACE 380
Db 361 SAFAQFAICLSSFETRIACE 380

RESULT 2
Q6YNP01_01 ARATH PRELIMINARY; PRT; 389 AA.
AC Q6YNP01_01 ARATH PRELIMINARY; PRT; 389 AA.
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DT 10-MAY-2005 (TRENBLrel. 30, Last annotation update)
DE At5g18680.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Chouk R., Chen H., Kim C.J., Shinn P., Ecker J.R.;
RA Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Kim C.J., Chen H., Chouk R., Shinn P., Ecker J.R.;
RA "Arabidopsis ORF clones.";
RT Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; BT010768; AAR23738.1; -; mRNA.
DR EMBL; BT020601; AAW80874.1; -; mRNA.
DR GO; GO:0006512; P-ubiquitin cycle; IEA.
DR InterPro; IPR001810; F-box.
DR InterPro; IPR000007; Tubby.
DR Pfam; PF00646; F-box; 1.
DR PRINTS; PR01573; SUPERTUBBY.
DR PROSITE; PS01200; TUB 1; UNKNOWN 1.
SQ SEQUENCE 389 AA; 43050 MW; 1270D94D2FB4E135 CRC64;

Query Match 73.5%; Score 1442.5; DB 2; Length 389;
Best Local Similarity 72.7%; Pred. No. 5.7e-119;
Matches 288; Conservative 32; Mismatches 53; Indels 23; Gaps 5;

QY 1 MTFRSILQEMRSRPHRVH-----AAASTANSDDPFSWSLPEELLREILIRVETVDCG 54
Db 1 MTLRSILQEMRSRPHRVHDLAAAAADSTSVSQDYRWSEIPEELLREILIRVEAADGG 60
QY 55 DWPSRRNVACAGVCSWRILTKIIVAVPEFSSKLTFFPISLKQSGPRDSLVCQFIKRNRN 114
Db 61 GWPSRRNVACAGVCSWRLLMNETVVVPEISKLTFPISLKQSGPRDSLVCQFIKRNRI 120
QY 115 TQSYHLYLGLTSLTDNGKFLAASKLKRATCTDYIISLRSDDISKRSNAYLGRMRSNFI 174
Db 121 TQSYHLYLGLTSLTDNGKFLAACKLKHHTCTDYIISLRSDDMRSRSQAYVGVKVSNE 180
QY 175 GTKFTVFDGS---QTGAAMOKRSNFIKVSPPVQSGPYPIAHISYELNVLGSRGPRM 231
Db 181 GTKFTVFDGNLSTGAALKRSRNPYPAKVSAPVPLGSPVAHITVELNVLGSRGPRM 240
QY 232 RCIMDTIPMSIVESRGVVASTSISSFSRSSPVFRSHKPLRSNSASCDSGNNLGDPL 291
Db 232 RCIMDTIPMSIVESRGVVASTSISSFSRSSPVFRSHKPLRSNSASCDSGNNLGDPL 291
```

```
Db 241 QCLMDTIPSTMEPOGVASEPSEFFLLGTRSTLSRSQSKPLRSS-----SHLKETPL 293
QY 292 VLSNKAPEWHQRLRCWCLNFRGRVTVASVKQFQVAV-SDCEAG-----QTSERILIQF 344
Db 294 VLSNKTPTWHQRLRCWCLNFRGRVTVASVKQFQVAV-SDCEAG-----QTSERILIQF 353
QY 345 GKVGKDMFTMDYGYPISAFOAFAICLSSFETRIACE 380
Db 354 GKVGKDMFTMDYGYPISAFOAFAICLSSFETRIACE 389

RESULT 3
Q6YNX7_01 ARATH PRELIMINARY; PRT; 380 AA.
AC Q6YNX7_01 ARATH PRELIMINARY; PRT; 380 AA.
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
DE Tubby-like protein 12.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Lai C.P., Lee C.L., Chen P.H., Wu S.H., Yang C.C., Shaw J.F.;
RA "Molecular analyses of the Arabidopsis TUBBY-like protein gene
RT family.";
RL Plant Physiol. 134:1586-1597(2004).
DR EMBL; AV046922; AAL03978.1; -; mRNA.
DR GO; GO:0006512; P-ubiquitin cycle; IEA.
DR InterPro; IPR001810; F-box.
DR InterPro; IPR000007; Tubby.
DR Pfam; PF00646; F-box; 1.
DR Pfam; PF01167; Tub; 1.
DR PRINTS; PR01573; SUPERTUBBY.
DR PROSITE; PS01200; TUB 1; UNKNOWN 1.
SQ SEQUENCE 380 AA; 41993 MW; EF52AA224CDA8F4D CRC64;

Query Match 72.1%; Score 1414.5; DB 2; Length 380;
Best Local Similarity 72.9%; Pred. No. 1.7e-116;
Matches 282; Conservative 31; Mismatches 51; Indels 23; Gaps 5;

QY 10 MRSRPHRVH-----AAASTANSDDPFSWSLPEELLREILIRVETVDCGWPSPRVV 63
Db 1 MRSRPHRVHDLAAAAADSTSVSQDYRWSEIPEELLREILIRVEAADGGWPSPRVV 60
QY 64 ACAGVCSWRILTKIIVAVPEFSSKLTFFPISLKQSGPRDSLVCQFIKRNNTQSYHLYLG 123
Db 61 ACAGVCSWRLLMNETVVVPEISKLTFPISLKQSGPRDSLVCQFIKRNITQSYHLYLG 120
QY 124 LTSLTDNGKFLAASKLKRATCTDYIISLRSDDISKRSNAYLGRMRSNFIQTKFTVFDG 183
Db 121 LTSLTDNGKFLAACKLKHHTCTDYIISLRSDDMRSRSQAYVGVKVSNEFLGKFTVFDG 180
QY 184 S---QTGAAMOKRSNFIKVSPPVQSGPYPIAHISYELNVLGSRGPRMRCIMDTIPM 240
Db 181 NLLPSTGAALKRSRNPYPAKVSAPVPLGSPVAHITVELNVLGSRGPRMRCIMDTIPT 240
QY 241 SVESRGVVASTSISSFSRSSPVFRSHKPLRSNSASCDSGNNLGDPLVLNKAAPRW 300
Db 241 STMEPOGVASEPSEFFLLGTRSTLSRSQSKPLRSS-----SHLKETPLVLNKTPTW 293
QY 301 HEQLRCWCLNFRGRVTVASVKQFQVAV-SDCEAG-----QTSERILIQFGKVGKDMFT 353
Db 294 HEQLRCWCLNFRGRVTVASVKQFQVAV-SDCEAG-----QTSERILIQFGKVGKDMFT 353
QY 354 MDYGYPISAFOAFAICLSSFETRIACE 380
Db 354 MDYGYPISAFOAFAICLSSFETRIACE 380
```

RESULT 4					
ID	Q8VY21_ ARATH PRELIMINARY;	PRT;	406 AA.		
AC	Q8VY21;				
DT	01-MAR-2002 (TrEMBLrel. 20, Created)				
DT	01-MAR-2002 (TrEMBLrel. 20, Last sequence update)				
DT	01-FEB-2005 (TrEMBLrel. 29, Last annotation update)				
DE	Tubby-like protein 3.				
GN	Name=At2g47900;				
OS	Arabidopsis thaliana (Mouse-ear cress).				
OC	Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;				
OC	rosids; eurosoids II; Brassicales; Brassicaceae; Arabidopsi.s.				
OX	NCBI_TaxID=3702;				
RN	[1]				
RP	NUCLEOTIDE SEQUENCE.				
RA	Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,				
RA	Goldsmith A.D., Lee J.M., Quach H.C., Toriumi M., Yu G., Bowser L.,				
RA	Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,				
RA	Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,				
RA	Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,				
RA	Seki M., Shinn P., Southwick A., Shinzaki K., Davis R.W., Ecker J.R.,				
RA	Theologis A.;				
RL	Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.				
RZ	[2]				
RP	NUCLEOTIDE SEQUENCE.				
RX	PubMed=15064372; DOI=10.1104/pp.103.037820;				
LA	Lai C.P., Lee C.L., Chen P.H., Wu S.H., Yang C.C., Shaw J.F.;				
RT	"Molecular analyses of the Arabidopsis TUBBY-like protein gene family";				
RT	Plant Physiol. 134:1586-1597(2004).				
RN	[3]				
RP	NUCLEOTIDE SEQUENCE.				
RA	Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,				
RA	Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,				
RA	Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,				
RA	Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,				
RA	Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,				
RA	Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,				
RA	Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinzaki K.,				
RA	Davis R.W., Ecker J.R., Theologis A.;				
RL	Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.				
DR	EMBL; AY074273; AAL66970.1; -; mRNA.				
DR	EMBL; AY045774; AAK98802.1; -; mRNA.				
DR	EMBL; AY096604; AAM20254.1; -; mRNA.				
DR	HSSP; P50586; I17E.				
DR	GO; GO:0045449; P:regulation of transcription; TAS.				
DR	InterPro; IPR001810; F-box.				
DR	InterPro; IPR000007; Tubby.				
DR	Pfam; PF00646; F-box; 1.				
DR	Pfam; PF01167; Tub; 1.				
DR	PRINTS; PR01573; SUPERTUBBY.				
DR	SEQUENCE 406 AA; 45311 MW; C9A8188C0AD077D4 CRC64;				
Qy	Query Match	58.6%;	Score 1149;	DB 2;	Length 406;
Bd	Best Local Similarity	56.7%;	Pred. No. 6.1e-93;		
Db	Matches 233;	Conservative 58;	Mismatches 84;	Indels 36;	Gaps 9
Qy	1 MTFPSLLQEM-----RSRPHRVHAAASTANSDFFSWSSELPPELL 41	:	:	:	:
Db	1 MSPKSLIQDMRGELGSI SRKGFDVRFGVGSRSRQVQDTSVPVDFAFKQSCWASMPPELL 60	:	:	:	:
Qy	42 REILIRVETVDGGDPSPRNVVACAGVCRSWRILTKEIVAVPEFSKLTFFPLSKSGPR 101	:	:	:	:
Db	61 RDVLMLRIEQSE-DTWPPSRKNVVSCAGVCGRNNRIVEIKVEI RVPELSKLTFFPLSKQGPGR 119	:	:	:	:
Qy	102 DSLVQCFTKRNNTQSXYHLGLTTLSTLDNGKFLLAASKLKRACTDYIIISLRSDISK 160	:	:	:	:
Db	120 GSLVQCYIMNRNSQTYLYLGLQAASNDGKFLAAKRFRPTCTDYIIISLCDVSR 179	:	:	:	:
Qy	161 RSNYALGRMRSNF LGTKPTVFDSQTG-AAKWQCSRSSNFI---KVSRPVQGSYPPIAH 216	:	:	:	:

180 GSNYYIGKLRNFIQTKFTYVDQAQTPNGTQVTRRSRLLSLKQVSPRIPSGNYPVAHI 239

217 SYELNVLGSRGPRRMRCIMDTIPMSIYESRGV-----VASTSISSFSSSSSPVFRSHK 270

240 SYELNVLGSRGPRRMQCWMDAIPASAVEPGCTAPTQTELHVNLSDSPFS--FFR--SK 295

271 PLRSNSASCDSDGNLGDPLVLNKNKAPRWHQOLRCWCLPHGRTVTVASVKNFOLVAV-S 329

296 SIRAEPSLPSGSSAAQKEGLVLNKNKAPRWHQOQCWCLNPNGRVTVTVASVKNFOLVAPE 355

330 DCEAGQTSERIIIOFGVKGDMFTMDYGYPISAFQAFAICLSSPETRIACE 380

356 NGPAGPEHENVILQFGVKGDVFTMDYQYPISAFQAFTICLSSPDTKIACE 406

RESULT 5

OB2257 ARATH

ID OB2257 ARATH PRELIMINARY; PRT; 407 AA.

AC OB2257;

DT 01-NOV-1998 (TREMBLrel. 08, Created)

DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)

DT 01-FEB-2005 (TREMBLrel. 29, Last annotation update)

DE Putative tubby protein.

GN Name=At2g47900;

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

OC rosids; eurosids I; Brassicales; Brassicaceae; Arabidopsi.

OC NCBI_Taxid=3702;

[1]

RP NUCLEOTIDE SEQUENCE.

RA Rounsley S.D., Lin X., Kaul S., Shea T.P., Fujii C.Y., Mason T.M.,

RA Shen M., Ronning C.M., Fraser C.M., Somerville C.R., Venter J.C.;

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

[2]

RP NUCLEOTIDE SEQUENCE.

RA Town C.D., Kaul S.;

RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.

[3]

RP NUCLEOTIDE SEQUENCE.

RA Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., VanAken S.E.,

RA Barnstead M.E., Mason T.M., Bowman C.L., Ronning C.M., Benito M.-I.,

RA Carstead A.J., Creasy T.H., Buell C.R., Town C.D., Nierman W.C.,

RA Fraser C.M., Venter J.C.;

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AC005309; AAC63644.1; -; Genomic DNA.

DR EMBL; AC006072; AA015124.1; -; Genomic_DNA.

DR PIR; H84920; H84920.

DR HSSP; P50586; 117E.

DR GO; GO:0006512; P:ubiquitin cycle; IEA.

DR InterPro; IPR001810; F-box.

DR InterPro; IPR000007; Tubby.

DR Pfam; PF00646; F-box; 1.

DR Pfam; PF01167; Tub; 1.

DR PRINTS; PR01573; SUPERTUBBY.

SQ SEQUENCE 407 AA; 45382 MW; 45AA55907F09DE86 CRC64;

Query Match 58.5%; Score 1147.5; DB 2; Length 407;

Best Local Similarity 56.6%; Pred. No. 8.3e-93;

Matches 233; Conservative 57; Mismatches 85; Indels 37; Gaps 9;

QY 1 MTFESLLQEM-----SSRPHRVVHAASTANSDDPFSWSELPELL 41

DB 1 MSFQSLIQDMRGELGSLRGKGFVRFQYGRSQRQVVQDTSVFDVAFQSCWASMPPELL 60

QY REILIRIVETVDGGDWPGRNRNVACAGYCRSNRILTKETIVAVPEPSSKLTTPISLKQSGPR 101

DB RDVLMRIEQSE-DTWPSRKNVSCAGVCRNWRREIVKSVIRVPELSKLTTPISLKQPGPR 119

QY DSLVQCFCKNRNNTQSHVLYGL--TTSLLTDNGKFLLAASKLKAATCTDIYIISLRSDIS 159

DB GSLVQCYIMRNRNOTYLYGLNQAAASNDGKFLLAAKFRFRPTCTDIYIISLNCDDVS 179

QY	160	KRSNAYLGRMRNFLTGTFTVDSQTG-AAKMQKRSRNF	1--KVSPRVQGSYPYIAH	215
Db	180	RGSNYVIGKLRNFLTGTFTVDAQFTNCTQVTRRSRLSLKQVSRIPSGNVPVAH	239	
QY	216	ISYELNVLSGRGPRRMRCIMDTIPMSIVSRGV-----VASTSISSPSRSPVPRSHS	269	
Db	240	ISYELNVLSGRGPRRMRCIMDTIPMSIVSRGV-----VASTSISSPSRSPVPRSHS	295	
QY	270	KPLRSNSASCSDSGNLGPPLVLSNKAPRWHEQLRCWCLNFHGRVTVASVKNFQLVAV-	328	
Db	296	KSIRAESLPSGSSAAQKGLLVLSNKAPRWHEQLRCWCLNFHGRVTVASVKNFQLVAV	355	
QY	329	SCEAAGQTSERIILQFGKVKDMFTMDYGYPISAFOAFAICLSSFPETRIACE	380	
Db	356	ENGAPGPEHNVILQFGKVKDMFTMDYGYPISAFOAFAICLSSFPETRIACE	407	
RESULT 6				
ID	Q8H485	ORYSA PRELIMINARY;	PRT;	406 AA.
AC	Q8H485			
DT	01-MAR-2003	(TREMBLrel. 23, Created)		
DT	01-MAR-2003	(TREMBLrel. 23, Last sequence update)		
DT	01-MAR-2004	(TREMBLrel. 26, Last annotation update)		
DE	Putative tubby related protein.			
GN	Name=P0450A04.125;			
OS	Oryza sativa (japonica cultivar-group).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;			
OC	Ehrhartoideae; Oryzeae; Oryza.			
OX	NCBI_TaxID=39947;			
RN	[1]			
RP	NUCLEOTIDE SEQUENCE.			
RA	Sasaki T., Matsumoto T., Yamamoto K.;			
RT	"Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, PAC			
RT	clone:P0450A04."			
RL	Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AP004274; BAC20077.1; -; Genomic_DNA.			
DR	HSSP; P50586; 117E.			
DR	Gramene; Q8H485; -.			
DR	GO; GO:0006512; P:ubiquitin cycle; IEA.			
DR	InterPro; IPR001810; F-box.			
DR	Pfam; PF00646; F-box; 1.			
DR	Pfam; PF01167; Tub; 1.			
DR	PRINTS; PR01573; SUPERTUBBY.			
DR	PROSITE; PS01200; TUB_1; UNKNOWN 1.			
DR	PROSITE; PS01201; TUB_2; UNKNOWN 1.			
SQ	SEQUENCE 406 AA; 45454 MW; 2FDIF9AE24FE8A4 CRC64;			
Query Match 58.1%; Score 1139; DB 2; Length 406;				
Best Local Similarity 56.4%; Pred. No. 4.7e-92;				
Matches 232; Conservative 57; Mismatches 86; Indels 36; Gaps 8;				
QY	1	MTFRSLQSEM-----RSRHRVVHAAASTANSDDPS	-----WSELPE 38	
Db	1	MSFRSVIQKVGKIGAISRRGRFRGRVRRVAAAEPPDESSAAALVWRSCVTQLPP	60	
QY	39	ELLREILIRVETVDGDPWFSRRNVACAGYCRSWRLITKEIVAVPFFSKLTPFISLKOS	98	
Db	61	ELLREVLARVESEGW-WPRRRDVACAGYCRSWRGIVREIVRTPEASGNLTFPISLKQP	119	
QY	99	GRDLSVQCFIKRNRNTQSHLYLGLTSLTDNGKFLAASKLRATCTDYIISLRSDI	158	
Db	120	GRDAPMKCFIVRNRTTQTYLYLGLTDALTDGKFLAARKRRTCTEYILISLMDMI	179	
QY	159	SKRSNAYLGRMRNFLTGTFTVDSQTG---AKMQKRSRNF	KVSPRVQGSYPYIA 214	
Db	180	SKRTDSYVGLKLRNFLTGTFTYIDAHPPYAGDVISKGQARVIGSNHLSFRPAGNYPVS	239	
QY	215	HISYELNVLSGRGPRRMRCIMDTIPMSIVSRGVASTSISSFSRSS	-----PVFRSHSK 270	
Db	240	HISYELNVLSGRGPRRMRCIMDTIPMSIVSRGVASTSISSFSRSS	-----PVFRSHSK 297	

QY	271	PLRSNSASCSDSGNLGPPLVLSNKAPRWHEQLRCWCLNFHGRVTVASVKNFQLVAVSD	330	
Db	298	SVRANNTASLITQN--GSKLVLSNKSPRWHEQLRCWCLNFHGRVTVASVKNFQLVASDE	355	
QY	331	CE-AQGTSERIILQFGKVKDMFTMDYGYPISAFOAFAICLSSFPETRIACE	380	
Db	356	SNPTNQEHDDVILQFGKVKDMFTMDYGYPISAFOAFAICLSSFPETRIACE	406	
RESULT 7				
ID	Q69U54	ORYSA PRELIMINARY;	PRT;	451 AA.
AC	Q69U54			
DT	25-OCT-2004	(TREMBLrel. 28, Created)		
DT	25-OCT-2004	(TREMBLrel. 28, Last sequence update)		
DT	25-OCT-2004	(TREMBLrel. 28, Last annotation update)		
DE	Putative chain A, C-terminal domain of mouse brain tubby protein.			
GN	Name=P0015C07.29;			
OS	Oryza sativa (japonica cultivar-group).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;			
OC	Ehrhartoideae; Oryzeae; Oryza.			
OX	NCBI_TaxID=39947;			
RN	[1]			
RP	NUCLEOTIDE SEQUENCE.			
RA	Sasaki T., Matsumoto T., Yamamoto K.;			
RT	"Oryza sativa nipponbare (GA3) genomic DNA, chromosome 8, PAC			
RT	clone:P0015C07."			
RL	Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AP004654; BAD33172.1; -; Genomic_DNA.			
DR	Gramene; Q69U54; -.			
DR	GO; GO:0006512; P:ubiquitin cycle; IEA.			
DR	InterPro; IPR001810; F-box.			
DR	InterPro; IPR000007; Tubby.			
DR	Pfam; PF00646; F-box; 1.			
DR	Pfam; PF01167; Tub; 1.			
DR	PRINTS; PR01573; SUPERTUBBY.			
DR	PROSITE; PS01200; TUB_1; UNKNOWN 1.			
DR	PROSITE; PS01201; TUB_2; 1.			
SQ	SEQUENCE 451 AA; 50528 MW; F875D1PB652B881E CRC64;			
Query Match 51.7%; Score 1015; DB 2; Length 451;				
Best Local Similarity 49.8%; Pred. No. 5.2e-81;				
Matches 212; Conservative 57; Mismatches 87; Indels 70; Gaps 9;				
QY	7	LOEMRSRPHRVVHAAASTANSDDPSFWSBELPEELREILIRVETVDGDPWFSRRNVACA	66	
Db	44	LSELDRPWW-----DQSRWVGLPELLRDVNMKLEEGE-SNWPSSKDVVACA	91	
QY	67	GVCRSWRILITKEIVAVPFFSKLTPFISLKQSGPRDSLVOCFIKRNRNTQSHLYLGLTT	126	
Db	92	AVCRTWREICKDIVOSPEICGKLTFFVSLKQPGPRDGLIQCFIKRDKSKLTLYLYLCLSP	151	
QY	127	S-LTDNGKFLAASKLRATCTDYIISLRSDIISKRSNAYLGRMRNFLTGTFTVED---	182	
Db	152	AVLSENGKFLAASKLRATCTDYIISVDKSNISRSSNGYGVCKMRNFLTGTFTVEDTQP	211	
QY	183	-----QTGAAMQKRSRNF	KVSPRVQGSYPYIAHISYELNVLSGRGPRRMRCI 234	
Db	212	PYNAGLSMCSQHSRRISRR-----VSPKLTGSGYPIAHVKYELNVLTGRPRMQCT	265	
QY	235	MDTIPMSIVSRGVASTSISSFSRSPVFRSHSKPLR-----SNSASCDS	282	
Db	266	MHSIFASAVDPGVPVGOPEQLLPGFESFESTNTSSRFSFMDRSLDFSSSRFSISGS	325	
QY	283	GNLGDPE-----PLVLSNKAPRWHEQLRCWCLNFHGRVTVASVKNFQLVAVSCEAG	334	
Db	326	ANQOGEDDIPEAKRPLVLNRKVPRWHEQLRCWCLNFHGRVTVASVKNFQLVAVSCEAG	385	
QY	335	QTSE-----RILOFGKVKDMFTMDYGYPISAFOAFAICLSSFE	374	
Db	386	QLEEQOQQOQQNHASSSSASDHGKVILOFGKVKDMFTMDYRYPLSAFOAFAICLTSFD	445	


```

QY 375 TRIACE 380
Db 446 TKLACE 451

RESULT 8
Q75HX5 ORYSA
ID Q75HX5 ORYSA PRELIMINARY; PRT; 445 AA.
AC Q75HX5;
DT 05-JUL-2004 (TREMELrel. 27, Created)
DT 05-JUL-2004 (TREMELrel. 27, Last sequence update)
DT 01-FEB-2005 (TREMELrel. 29, Last annotation update)
DE Putative tubby protein.
GN Name=OSJNB0042J17.6; Synonyms=OJ1058_F05.11;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoideae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Chow T.-Y., Hsing Y.-I.C., Chen C.-S., Chen H.-H., Liu S.-M.,
RA Chao Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R.,
RA Chen Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H.,
RA Hsiung J.-N., Hsu C.-H., Huang J.-J., Kau P.-I., Lee M.-C., Leu H.-L.,
RA Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,
RA Wu H.-P., Shaw J.-F.;
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Chow T.-Y., Hsing Y.-I.C., Chen C.-S., Chen H.-H., Liu S.-M.,
RA Chao Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R.,
RA Chen Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H.,
RA Hsiung J.-N., Hsu C.-H., Huang J.-J., Kau P.-I., Lee M.-C., Leu H.-L.,
RA Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,
RA Wu H.-P., Shaw J.-F.;
RT "Oryza sativa BAC OJ1058_F05 genomic sequence.";
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC134930; RAT07611.1; -; Genomic DNA.
DR EMBL; AC105318; RAV59313.1; -; Genomic DNA.
DR Gramene; Q75HX5; -;
DR GO; GO:0006512; Pubiquitin cycle; IEA.
DR InterPro; IPR001810; F-box.
DR InterPro; IPR000007; Tubby.
DR Pfam; PF00646; F-box; 1.
DR Pfam; PF01167; Tub; 1.
DR PRINTS; PR01573; SUPERTUBBY.
DR PROSITE; PS01200; TUB 1; UNKNOWN_1.
DR PROSITE; PS01201; TUB 2; 1.
SQ SEQUENCE 445 AA; 49554 MW; D3E1BF1AB11ADCF CRC64;

Query Match 51.7%; Score 1013.5; DB 2; Length 445;
Best Local Similarity 51.3%; Pred. No. 7e-81;
Matches 211; Conservative 66; Mismatches 91; Indels 43; Gaps 11;

QY 11 RSRHRVHAASASTANSDPFSWSELPELLREILIRVETVDGOWPSSRRNVACAGVCR 70
Db 37 RCKSQSTVHELCDADLIQESRWASLPPELLRDVIRLEAS-STWPSRKDVVSCAAVCK 95
QY 71 SWRLITKEIVAVPESSKLTFFISLKQSGPRSLVQCFTKRNNTQSYHLVGLTTS-IT 129
Db 96 AWRECKEIVLSPFCGKLTFFLSLKQPGPRGMIQCFIKRDKSKSYHLVLCLSFVAVLA 155
QY 130 DNGKLLAASKLRATCTDYIISLSDDIKSNAYLGHMRNFLTGTFTVD--GSQGTG 187
Db 156 DSGKFLLSAKRHKTKCTEYVISMADNISRSSSTYIGKLRNFLTGTFTFIYDTQPSYNG 215
QY 188 AAKMQKSRSS---NFIKSPRPQGSYPTAHSYELNVLGSGPRMRRCIMDTIPMSIVE 244
Db 216 AVIPVGRSSRRFNKSKVSPKPPSSGYNTAQVYELNVLGTRGRMRHCVMSIPASSVE 275
QY 245 SRGVV-----ASTSISFSRSSGPFVRS--HSKPLRSNSASCSD-----S 282

```

```

Db 276 FGIIVPGQDEQIVPRAPESFRSTTSF-SKSSIMDSMDFFSSSRDFSDIAGGTIN 334
QY 283 GNNLG---DPLVLNKNAPRWHEQLRCWCLNPHGRVTVASVKNFQLVAV-SDEAGOTS- 337
Db 335 GDEEGNKERPLVLNKNAPRWHEQLRCWCLNPHGRVTVASVKNFQLVAV-SDEAGOTS- 394
QY 338 -----ERITLQSKVKCKDMFTWDYGYPISAFQAFALCLSSFETRIACE 380
Db 395 SQAPPEQDKITLQFGKVAQDMFTWDYRYPLSAFQAFALCLSSFDTKLACE 445

RESULT 9
Q9ZP59 ARATH
ID Q9ZP59 ARATH PRELIMINARY; PRT; 455 AA.
AC Q9ZP59;
DT 01-MAY-1999 (TREMELrel. 10, Created)
DT 01-MAY-1999 (TREMELrel. 10, Last sequence update)
DT 01-FEB-2005 (TREMELrel. 29, Last annotation update)
DE F22K20.1 protein (Tub family protein, putative) (Atlg76900/F7012_7)
DE (Tubby-like protein TULP1).
GN Name=F22K20.1; Synonyms=F7012.7, TULP1;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Federspiel N.A., Palm C.-J., Conway A.B., Kurtz D.B., Conway A.R.,
RA Au M., Araujo R., Buehler E., Dewar K., Feng J., Kim C., Li Y.,
RA Oji O., Osborne B.I., Shinn P., Sun H., Toriumi M., Vysotskaia V.S.,
RA Yu G., Ecker J., Theologis A., Davis R.W.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,
RA Mafti R., Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Fraser C.M.;
RA Barnstead M.E., Bowman C.L., White O., Niernan W.C., Fraser C.M.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Shinn P., Chen H., Cheuk R., Kim C.J., Bowser L., Carninci P.,
RA Chan M.M., Chang C.H., Dale J.M., Deng J.M., Hayashizaki Y.,
RA Hsuan V.W., Iehida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M.,
RA Southwick A., Tang C.C., Toriumi M., Wallender E.K., Wong C., Wu H.C.,
RA Yamada K., Yu G., Yuan S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE.
RA Cheuk R., Chen H., Kim C.J., Shinn P., Bowser L., Carninci P.,
RA Chan M.M., Chang C.H., Dale J.M., Hayashizaki Y., Hsuan V.W.,
RA Iehida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Wong C., Wu H.C., Yamada K., Yu G., Yuan S.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
RN [5]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15064372; DOI=10.1104/pp.103.037820;
RA Lai C.P., Lee C.L., Chen P.H., Wu S.H., Yang C.C., Shaw J.F.;
RT "Molecular analyses of the Arabidopsis TUBBY-like protein gene
family.";
RL Plant Physiol. 134:1586-1597(2004).
DR EMBL; AC002291; AAC00626.1; -; Genomic DNA.
DR EMBL; AC079283; AAG51146.1; -; Genomic DNA.
DR EMBL; AY139758; RAN98079.1; -; mRNA.
DR EMBL; BT003039; AAO23604.1; -; mRNA.
DR EMBL; AF487267; AAQ06240.1; -; mRNA.

```



```
AC Q9ST78;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE CAA303719.1 protein.
GN Name=q3037.19;
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]_
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=DNA.
RA Hong G., Chen Z.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ245900; CAB53492.1; -; Genomic DNA.
DR HSSP; P50586; 1176.
DR Gramene; O9ST78; -.
DR GO; GO:0006512; Ubiquitin cycle; IEA.
DR InterPro; IPR001810; F-box.
DR InterPro; IPR000007; Tubby.
DR Pfam; PF00646; F-box; 1.
DR Pfam; PF01167; Tub; 1.
DR PRINTS; PR01573; SUPERTUBBY.
DR PROSITE; PS01200; TUB_1; UNKNOWN_1.
DR PROSITE; PS01201; TUB_2; 1.
DR SEQUENCE 462 AA; 51513 MW; 456367D114AE59F4 CRC64;

Query Match 50.4%; Score 989; DB 2; Length 462;
Best Local Similarity 44.8%; Pred. No. 1.le-78;
Matches 210; Conservative 70; Mismatches 93; Indels 96; Gaps 10;

QY 1 MTFRSLLOEMR-----SRPVRVHAAASTANSSDPF-----SNSELP 37
DB 1 MSFRSIVRDRDGFGLSGRRGFEVRAILGRRKRGSHGAVHELHDPVPVIOSSCWASLPPEL 60
QY 38 EELLREILRVETVGGDWPSPRRNVACAGVCSRWRIITKEIVAVPEFSKLTFFISLKQ 97
DB 61 PELLRDVVKRLEE--DSDNWPSPKOVVACASVCTTWRDMCKDVRNPEFCGKLTFFVSLQ 119
QY 98 SGRPSLVCQCFKRNNTQSYHLGLTTS-LTNGKFLAASKLKRATCTDYIISLRSD 156
DB 120 PPRDGVIOCFIKRDKSKLTYHLCLSSAVLDGKFLSKRSPRTHTDYIISMDSK 179
QY 157 DISKSNAYLGRMSNFLTGTFTVPD-----GSQTGAAMQKSRSSNFTKVSFVPQGS 211
DB 180 NISRSSSGYIGKLSNFLTGTFTFYDTPPNARTLCSQERTSRRPSSRKVPKVTGCV 239
QY 212 PTAHSYELNVLGSRGPRRMRCIMDTPMSIVESRGVV-----ASTS 253
DB 240 PIVQVNYELNVLGTRGPRMQCAMHSIPASAVEPGGIVFGQPKELLPLRFESFRSMAT 299
QY 254 ISFSRSRSPVFRSHSKPLRSNSASCSDGNNLGL-----DCEAGQTSER----- 339
DB 300 FKYS-----ITDHTSDSSRPSFEGGALQCGEQEDGDVVKERPLVLRNKAAPRW 353
QY 302 EQLRCWCLNFHGRVTVASVKNFQLVAVS-----DCEAGQTSER----- 380
DB 354 EQLQCLNFRGRVTVASVKNFQLIAAQPASGAASEPSEQAAQOQQTQPSQSSSS 413
QY 340 -----IILQFGVKGMFTMDYGPISAFQFAICLSSPETRIACE 380
DB 414 SSSSNHDTVILQFGKVAKMFTMDYRPLSAFOFAICLTSTFDTKLACE 462

RESULT 14
Q53PP5_ORYSA
ID Q53PP5_ORYSA PRELIMINARY; PRT; 440 AA.
AC Q53PP5;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)

DE Tub family, putative.
GN ORFNames=LOC_Os11g06420;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]_
RP NUCLEOTIDE SEQUENCE.
RA Buell C., Yuan Q., Ouyang S., Liu J., Wang A., Maiti R., Lin H.,
RA Zhu W., Hamilton J., Jones K., Tallon L., Feldblyum T., Teitlin T.,
RA Bera J., Kim M., Jin S., Fadrosch D., Vuong H., Overton II L.,
RA Reardon M., Weaver B., Johri S., Lewis M., Utterback T., Van Aken S.,
RA Wortman J., Haas B., Koo H., Zismann V., Hsiao J., Iobst S.,
RA de Vazeilles A., White O., Salzberg S., Fraser C.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [2]_
RP NUCLEOTIDE SEQUENCE.
RA Buell R.;
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC120533; AAX95106.1; -; Genomic DNA.
SQ SEQUENCE 440 AA; 48955 MW; 9D9D7312F67AB9B4 CRC64;

Query Match 49.4%; Score 969; DB 2; Length 440;
Best Local Similarity 46.2%; Pred. No. 6e-77;
Matches 207; Conservative 72; Mismatches 93; Indels 76; Gaps 12;

QY 1 MTFRSLLOEM-----RSRPRVHAAASTANSSDPFWSSELPEEL 40
DB 1 MSFRSIVRDRDGFGLSGRRGFEVRAILGRRKRGSHGAVHELHDPVPVIOSSCWASLPPEL 60
QY 41 LREILRVETVGGDWPSPRRNVACAGVCSRWRIITKEIVAVPEFSKLTFFISLKQSGP 100
DB 61 LRDIITLERLEEE--ATWPSRKHVACAGVCTWRCKEIVKNPELGGKITFFISLRQGP 119
QY 101 RDSLVCQCFKRNNTQSYHLGLTTS-LTNGKFLAASKLKRATCTDYIISLRSDDIS 159
DB 120 RDGTWQCIFIRDKSTQTYLYLSLGSVLDVNGKFLSKRNWHATCTEYVISMANNLS 179
QY 160 KRSNAYLGRMSNFLTGTFTVPD-----GSQTGAAMQKSRSSNFTKVSFVPQGS 210
DB 180 RSTNTNIGKLSNFLTGTFTFYDTPYNATSDSQSGKT---SRRFSN--KGTAKHPCST 234
QY 211 YPIAHSYELNVLGSRGPRRMRCIMDTPMSIVESRGVVAST-----SISFSRSRSP 263
DB 235 YSIANISYELNVLGTRGPRRMCMCLMHSIPASSLEAGTVPSQPDILAHLSNESSFRSV 294
QY 264 VPRS-----HSKPLRSNSASCSD-----SGNNLG-----DPLVLNKAAPRWHEQL 304
DB 295 FSKSSVMDHS--MHFSSAQFSDISIGDPRIGRVLSDDECKETPLILQNKAPRWHEQL 352
QY 305 RCWCLNFHGRVTVASVKNFQLVAVSDCEAGQTS-----ERILQFGVKGMF 352
DB 353 QCWCLNFRGRVTVASVKNFQLIAATQPAAGAPTSPQVPPPPPHDKVILQFGKVAKMDF 412
QY 353 TMDYGVPIASAFQFAICLSSPETRIACE 380
DB 413 TMDYHPLSAFOFAICLSSFDTKLACE 440

RESULT 15
Q8GVES_ARATH
ID Q8GVES_ARATH PRELIMINARY; PRT; 394 AA.
AC Q8GVES;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Tubby-like protein 2.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
```

```
[1]
RN NUCLEOTIDE SEQUENCE.
RX PubMed=15064372; DOI=10.1104/pp.103.037820;
Lai C.P., Lee C.L., Chen P.H., Wu S.H., Yang C.C., Shaw J.F.;
RT "Molecular analyses of the Arabidopsis TUBBY-like protein gene
family.";
RL Plant Physiol. 134:1586-1597(2004).
DR EMBL; AY045773; AAK98801.1; -; mRNA.
DR HSSP; PS0586; IIIE.
DR GO; GO:0045449; P:regulation of transcription; TAS.
DR InterPro; IPR001810; F-box.
DR InterPro; IPR000007; Tubby.
DR Pfam; PF00646; F-box; 1.
DR Pfam; PF01167; Tub; 1.
DR PRINTS; PR01573; SUPERTUBBY.
DR PROSITE; PS01200; TUB 1; UNKNOWN 1.
DR PROSITE; PS01201; TUB 2; UNKNOWN 1.
SQ SEQUENCE 394 AA; 43878 MW; 2BBA39F8666AEDC0 CRC64;

Query Match 49.3%; Score 966.5; DB 2; Length 394;
Best Local Similarity 50.1%; Pred. No. 8.5e-77;
Matches 201; Conservative 64; Mismatches 107; Indels 29; Gaps 7;

QY 1 MTFRSLQEMR-----SRPHRVHAAASTANSDPFS-WSELPEELREIL 45
Db | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 46 IRVETVGGDMFSPRRNVACAGVCRSWRLTKETVAVPFSSKLTPTPISLKQSGPRDSIV 105
Db ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
61 WRVESETA-WPARAAVVCASVCKSWRGITWEIVRIPEQCGKLTPTPISLKQSGPRDSPI 119
QY 106 QCFIKENRNTQSYHLGLTSLTNGKFLAASKLKRATCTDYIISLRDDISKRSNAY 165
Db ||||| ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
120 QCFIKENRATATYILYGLMPSETENDKLLAARRIRATCTDFIISLAKNFSRSSSTY 179
QY 166 LGRMSNFLGTFTFDG-SQTAAKQKSRSNFTKVPVPOGSPYIAHISYELNLVG 224
Db | : : ||| ||| : : : : : | : : : : : | : : : : : | : : : : : |
180 VKLSGFLGTFTYIDNQTASTAQAPNRLHPKQAPKLTNSSTVGNITYELNVL 239
QY 225 SRGPRMRCIMDTPMSIYVESRGVVASTSISSFSSRSPVFR--SHSKPLRSNSASCSDS 282
Db : ||||| : ||| : : : : : | : : : : : | : : : : : | : : : : : |
240 TRGPRMHCAMDSIPLSSVIAPSPVVGTEEEVSSSPSPKGETITTDKEIPDNSPS--- 295
QY 283 GNNLGDPPPLVLSKNAPRWEHLRCWCLNFHGRVTVASVKNFOLVAVSDCEAG----QTSE 339
Db ||||| ||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
296 ---LRDQPLVLKNKSPRWEHLQCWCLNFKGRVTVASVKNFOLVAVSDAPPEHER 352
QY 340 IILQFGKVGKDMFTMDYGYPISAFQAFACLSSFETRIACE 380
Db : ||||| : ||| : ||||| : ||||| : ||||| : ||||| : ||||| : |||
353 VILQFGKIGKIDFTMDYRPLSAFQAFACISSFDTKPACE 393
```

Search completed: December 23, 2005, 23:27:49
Job time : 235 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 23, 2005, 23:20:27 ; Search time 46 Seconds
(without alignments)
682.973 Million cell updates/sec

Title: US-10-763-042-9
Perfect score: 1962
Sequence: 1 MFRSLQLQMRSPRHRVHA.....SAFQAFATLSSFFETRIACE 380

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/iaa/5 COMB.pep.*
2: /cgn2_6/prodata/1/iaa/6 COMB.pep.*
3: /cgn2_6/prodata/1/iaa/H COMB.pep.*
4: /cgn2_6/prodata/1/iaa/FACTUS COMB.pep.*
5: /cgn2_6/prodata/1/iaa/RE COMB.pep.*
6: /cgn2_6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	441.5	22.5	506	2 US-09-032-365A-19	Sequence 19, Appl
2	437.5	22.3	460	1 US-08-630-592-7	Sequence 7, Appl
3	437.5	22.3	460	1 US-08-714-991-7	Sequence 7, Appl
4	437.5	22.3	460	2 US-09-032-365A-8	Sequence 8, Appl
5	437.5	22.3	506	1 US-08-631-200-8	Sequence 8, Appl
6	437.5	22.3	506	1 US-08-829-553-8	Sequence 8, Appl
7	437.5	22.3	506	1 US-08-922-267A-8	Sequence 8, Appl
8	437.5	22.3	506	1 US-08-936-707A-8	Sequence 8, Appl
9	437.5	22.3	506	1 US-08-936-706A-8	Sequence 8, Appl
10	437.5	22.3	506	2 US-09-248-203-8	Sequence 8, Appl
11	437.5	22.3	506	2 US-09-032-365A-62	Sequence 62, Appl
12	437.5	22.3	506	2 US-08-812-824-4	Sequence 4, Appl
13	437.5	22.3	506	2 US-09-406-071-8	Sequence 8, Appl
14	437.5	22.3	506	2 US-08-955-918C-10	Sequence 10, Appl
15	437.5	22.3	506	2 US-08-697-766A-10	Sequence 10, Appl
16	437.5	22.3	506	2 US-09-814-986-8	Sequence 8, Appl
17	437.5	22.3	512	2 US-09-032-365A-60	Sequence 60, Appl
18	437.5	22.3	518	1 US-09-032-365A-58	Sequence 58, Appl
19	437.5	22.3	561	1 US-08-714-991-27	Sequence 27, Appl
20	437.5	22.3	561	2 US-09-032-365A-10	Sequence 10, Appl
21	436.5	22.2	459	1 US-08-630-592-2	Sequence 2, Appl
22	436.5	22.2	459	1 US-08-714-991-2	Sequence 2, Appl
23	436.5	22.2	459	1 US-09-032-365A-2	Sequence 2, Appl
24	436.5	22.2	505	1 US-08-631-200-2	Sequence 2, Appl
25	436.5	22.2	505	1 US-08-630-592-4	Sequence 4, Appl
26	436.5	22.2	505	1 US-08-714-991-4	Sequence 4, Appl
27	436.5	22.2	505	1 US-08-829-553-2	Sequence 2, Appl

28	436.5	22.2	505	1 US-08-922-267A-2	Sequence 2, Appl
29	436.5	22.2	505	1 US-08-936-707A-2	Sequence 2, Appl
30	436.5	22.2	505	1 US-08-936-706A-2	Sequence 2, Appl
31	436.5	22.2	505	2 US-09-248-203-2	Sequence 2, Appl
32	436.5	22.2	505	2 US-09-032-365A-4	Sequence 4, Appl
33	436.5	22.2	505	2 US-08-812-824-3	Sequence 3, Appl
34	436.5	22.2	505	2 US-09-406-071-2	Sequence 2, Appl
35	436.5	22.2	505	2 US-08-914-986-2	Sequence 2, Appl
36	427	21.8	504	2 US-08-955-918C-7	Sequence 7, Appl
37	427	21.8	504	2 US-08-697-766A-7	Sequence 7, Appl
38	423	21.6	460	2 US-09-270-767-45069	Sequence 45069, A
39	418	21.3	285	1 US-08-631-200-15	Sequence 15, Appl
40	418	21.3	285	1 US-08-829-553-15	Sequence 15, Appl
41	418	21.3	285	1 US-08-922-267A-15	Sequence 15, Appl
42	418	21.3	285	1 US-08-936-707A-15	Sequence 15, Appl
43	418	21.3	285	1 US-08-936-706A-15	Sequence 15, Appl
44	418	21.3	285	2 US-09-248-203-15	Sequence 15, Appl
45	418	21.3	285	2 US-09-406-071-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1
US-09-032-365A-19
; Sequence 19, Application US/09032365A
; Patent No. 6114502
; GENERAL INFORMATION:
; APPLICANT: No. 6114502th, Michael
; APPLICANT: Nishina, Patey
; APPLICANT: Naggart, Juergen
; APPLICANT: No. 6114502en-Trauth, Konrad
; TITLE OF INVENTION: GENE FAMILY ASSOCIATED WITH
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Bozicevic & Reed, LLP
; STREET: 285 Hamilton Avenue, Suite 200
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/032,365A
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: SEQ-2CIP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-327-3400
; TELEFAX: 650 327-3231
; TELEX:
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 506 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-032-365A-19
Query Match 22.5%; Score 441.5; DB 2; Length 506;
Best Local Similarity 36.9%; Pred. No. 1.8e-40;

Matches 106; Conservative 44; Mismatches 90; Indels 47; Gaps 7;	
QY 100	PRDSVOCIFKRNNTQ-----SYHLVGLTSLTDNGKFLAASKLRATCTDYILSL 153
DB 261	PRDHVMQWIRVRNKGMDKMPFSYYLYLEGEDGV---AHFLAGRRKRKSTSNYLSL 317
QY 154	RSDDISKRNAVILGRMRSNPLGTFKFTVFDGSGTGAARMQKSRSSNFVKVSPRPQGSYPI 213
DB 318	DPKDSRNGSNFVGKVRSNVLGTFKFTFDNG-----VNPERSYVWPDSAIR-----EL 367
QY 214	AHISYELNVLGRGPRMRCIMDTPMISVSRGVVASTSISFSRSPVFRSHSKPLR 273
DB 368	GVVCYETNVLGRGPRKMTVI-----LPGMDSRKQ---RMKVQPON 405
QY 274	SNSASCSDSGNLGDPLVLSKAPRWHLQRCWCLNFHGRVTVASVKNFOLVAVSDCEA 333
DB 406	QDSDILSRVQKGAGHGLLLQNKAPSWSDSGAYVLNFGHGRVTRASVKNFQIVHPDE--- 462
QY 334	GOTSERIILQFGKVGKDMFTMDYGYPISAFOAFAICLSSFETRIACE 380
DB 463	---PDLVLQGRVAPNIFTMDFRYPLCPLOAFAICLSSFDGKLACE 506
RESULT 2	
US-08-630-592-7	
; Sequence 7, Application US/08630592	
; Patent No. 5770432	
; GENERAL INFORMATION:	
; APPLICANT: Nishina, Patsy	
; APPLICANT: No. 5770432entTrauth, Konrad	
; APPLICANT: Naggett, Juergen	
; APPLICANT: No. 5770432th, Michael	
; TITLE OF INVENTION: Obesity Associated Genes	
; NUMBER OF SEQUENCES: 25	
; CORRESPONDENCE ADDRESS:	
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT	
; STREET: 3400 Embarcadero Center, Suite 3400	
; CITY: San Francisco	
; STATE: California	
; COUNTRY: USA	
; ZIP: 941114187	
; COMPUTER READABLE FORM:	
; MEDIUM TYPE: Floppy disk	
; COMPUTER: IBM PC compatible	
; OPERATING SYSTEM: PCDOS/MSDOS	
; SOFTWARE: Patent In Release #1.0, Version #1.30	
; CURRENT APPLICATION DATA:	
; APPLICATION NUMBER: US/08/630,592	
; FILING DATE:	
; CLASSIFICATION: 435	
; ATTORNEY/AGENT INFORMATION:	
; NAME: Sherwood, Pamela J.	
; REGISTRATION NUMBER: 36,677	
; REFERENCE/DOCKET NUMBER: A59504/BIR/PJS	
; TELECOMMUNICATION INFORMATION:	
; TELEPHONE: (415) 7811989	
; TELEFAX: (415) 3983249	
; TELEX: 910 277299	
; INFORMATION FOR SEQ ID NO: 7:	
; SEQUENCE CHARACTERISTICS:	
; LENGTH: 460 amino acids	
; TYPE: amino acid	
; STRANDEDNESS: double	
; TOPOLOGY: linear	
; MOLECULE TYPE: peptide	
US-08-630-592-7	
Query Match 22.3%; Score 437.5; DB 1; Length 460;	
Best Local Similarity 34.9%; Pred. No. 4.3e-40;	
Matches 109; Conservative 60; Mismatches 82; Indels 61; Gaps 12;	
QY 78	EIVAVPEFSSKLTFFPISLQSGPRDSLVOCFIKRNRNTQ-----SYHLVGLTSLTDN 131

Matches 106; Conservative 44; Mismatches 90; Indels 47; Gaps 7;	
QY 100	PRDSVOCIFKRNNTQ-----SYHLVGLTSLTDNGKFLAASKLRATCTDYILSL 153
DB 261	PRDHVMQWIRVRNKGMDKMPFSYYLYLEGEDGV---AHFLAGRRKRKSTSNYLSL 317
QY 154	RSDDISKRNAVILGRMRSNPLGTFKFTVFDGSGTGAARMQKSRSSNFVKVSPRPQGSYPI 213
DB 318	DPKDSRNGSNFVGKVRSNVLGTFKFTFDNG-----VNPERSYVWPDSAIR-----EL 367
QY 214	AHISYELNVLGRGPRMRCIMDTPMISVSRGVVASTSISFSRSPVFRSHSKPLR 273
DB 368	GVVCYETNVLGRGPRKMTVI-----LPGMDSRKQ---RMKVQPON 405
QY 274	SNSASCSDSGNLGDPLVLSKAPRWHLQRCWCLNFHGRVTVASVKNFOLVAVSDCEA 333
DB 406	QDSDILSRVQKGAGHGLLLQNKAPSWSDSGAYVLNFGHGRVTRASVKNFQIVHPDE--- 462
QY 334	GOTSERIILQFGKVGKDMFTMDYGYPISAFOAFAICLSSFETRIACE 380
DB 463	---PDLVLQGRVAPNIFTMDFRYPLCPLOAFAICLSSFDGKLACE 506
RESULT 2	
US-08-630-592-7	
; Sequence 7, Application US/08630592	
; Patent No. 5770432	
; GENERAL INFORMATION:	
; APPLICANT: Nishina, Patsy	
; APPLICANT: No. 5770432entTrauth, Konrad	
; APPLICANT: Naggett, Juergen	
; APPLICANT: No. 5770432th, Michael	
; TITLE OF INVENTION: Obesity Associated Genes	
; NUMBER OF SEQUENCES: 25	
; CORRESPONDENCE ADDRESS:	
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT	
; STREET: 3400 Embarcadero Center, Suite 3400	
; CITY: San Francisco	
; STATE: California	
; COUNTRY: USA	
; ZIP: 941114187	
; COMPUTER READABLE FORM:	
; MEDIUM TYPE: Floppy disk	
; COMPUTER: IBM PC compatible	
; OPERATING SYSTEM: PCDOS/MSDOS	
; SOFTWARE: Patent In Release #1.0, Version #1.30	
; CURRENT APPLICATION DATA:	
; APPLICATION NUMBER: US/08/630,592	
; FILING DATE:	
; CLASSIFICATION: 435	
; ATTORNEY/AGENT INFORMATION:	
; NAME: Sherwood, Pamela J.	
; REGISTRATION NUMBER: 36,677	
; REFERENCE/DOCKET NUMBER: A59504/BIR/PJS	
; TELECOMMUNICATION INFORMATION:	
; TELEPHONE: (415) 7811989	
; TELEFAX: (415) 3983249	
; TELEX: 910 277299	
; INFORMATION FOR SEQ ID NO: 7:	
; SEQUENCE CHARACTERISTICS:	
; LENGTH: 460 amino acids	
; TYPE: amino acid	
; STRANDEDNESS: double	
; TOPOLOGY: linear	
; MOLECULE TYPE: peptide	
US-08-630-592-7	
Query Match 22.3%; Score 437.5; DB 1; Length 460;	
Best Local Similarity 34.9%; Pred. No. 4.3e-40;	
Matches 109; Conservative 60; Mismatches 82; Indels 61; Gaps 12;	
QY 78	EIVAVPEFSSKLTFFPISLQSGPRDSLVOCFIKRNRNTQ-----SYHLVGLTSLTDN 131

Matches 106; Conservative 44; Mismatches 90; Indels 47; Gaps 7;	
QY 100	PRDSVOCIFKRNNTQ-----SYHLVGLTSLTDNGKFLAASKLRATCTDYILSL 153
DB 261	PRDHVMQWIRVRNKGMDKMPFSYYLYLEGEDGV---AHFLAGRRKRKSTSNYLSL 317
QY 154	RSDDISKRNAVILGRMRSNPLGTFKFTVFDGSGTGAARMQKSRSSNFVKVSPRPQGSYPI 213
DB 318	DPKDSRNGSNFVGKVRSNVLGTFKFTFDNG-----VNPERSYVWPDSAIR-----EL 367
QY 214	AHISYELNVLGRGPRMRCIMDTPMISVSRGVVASTSISFSRSPVFRSHSKPLR 273
DB 368	GVVCYETNVLGRGPRKMTVI-----LPGMDSRKQ---RMKVQPON 405
QY 274	SNSASCSDSGNLGDPLVLSKAPRWHLQRCWCLNFHGRVTVASVKNFOLVAVSDCEA 333
DB 406	QDSDILSRVQKGAGHGLLLQNKAPSWSDSGAYVLNFGHGRVTRASVKNFQIVHPDE--- 462
QY 334	GOTSERIILQFGKVGKDMFTMDYGYPISAFOAFAICLSSFETRIACE 380
DB 463	---PDLVLQGRVAPNIFTMDFRYPLCPLOAFAICLSSFDGKLACE 506
RESULT 2	
US-08-630-592-7	
; Sequence 7, Application US/08630592	
; Patent No. 5770432	
; GENERAL INFORMATION:	
; APPLICANT: Nishina, Patsy	
; APPLICANT: No. 5770432entTrauth, Konrad	
; APPLICANT: Naggett, Juergen	
; APPLICANT: No. 5770432th, Michael	
; TITLE OF INVENTION: Obesity Associated Genes	
; NUMBER OF SEQUENCES: 25	
; CORRESPONDENCE ADDRESS:	
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT	
; STREET: 3400 Embarcadero Center, Suite 3400	
; CITY: San Francisco	
; STATE: California	
; COUNTRY: USA	
; ZIP: 941114187	
; COMPUTER READABLE FORM:	
; MEDIUM TYPE: Floppy disk	
; COMPUTER: IBM PC compatible	
; OPERATING SYSTEM: PCDOS/MSDOS	
; SOFTWARE: Patent In Release #1.0, Version #1.30	
; CURRENT APPLICATION DATA:	
; APPLICATION NUMBER: US/08/630,592	
; FILING DATE:	
; CLASSIFICATION: 435	
; ATTORNEY/AGENT INFORMATION:	
; NAME: Sherwood, Pamela J.	
; REGISTRATION NUMBER: 36,677	
; REFERENCE/DOCKET NUMBER: A59504/BIR/PJS	
; TELECOMMUNICATION INFORMATION:	
; TELEPHONE: (415) 7811989	
; TELEFAX: (415) 3983249	
; TELEX: 910 277299	
; INFORMATION FOR SEQ ID NO: 7:	
; SEQUENCE CHARACTERISTICS:	
; LENGTH: 460 amino acids	
; TYPE: amino acid	
; STRANDEDNESS: double	
; TOPOLOGY: linear	
; MOLECULE TYPE: peptide	
US-08-630-592-7	
Query Match 22.3%; Score 437.5; DB 1; Length 460;	
Best Local Similarity 34.9%; Pred. No. 4.3e-40;	
Matches 109; Conservative 60; Mismatches 82; Indels 61; Gaps 12;	
QY 78	EIVAVPEFSSKLTFFPISLQSGPRDSLVOCFIKRNRNTQ-----SYHLVGLTSLTDN 131

Matches 106; Conservative 44; Mismatches 90; Indels 47; Gaps 7;	
QY 100	PRDSVOCIFKRNNTQ-----SYHLVGLTSLTDNGKFLAASKLRATCTDYILSL 153
DB 261	PRDHVMQWIRVRNKGMDKMPFSYYLYLEGEDGV---AHFLAGRRKRKSTSNYLSL 317
QY 154	RSDDISKRNAVILGRMRSNPLGTFKFTVFDGSGTGAARMQKSRSSNFVKVSPRPQGSYPI 213
DB 318	DPKDSRNGSNFVGKVRSNVLGTFKFTFDNG-----VNPERSYVWPDSAIR-----EL 367
QY 214	AHISYELNVLGRGPRMRCIMDTPMISVSRGVVASTSISFSRSPVFRSHSKPLR 273
DB 368	GVVCYETNVLGRGPRKMTVI-----LPGMDSRKQ---RMKVQPON 405
QY 274	SNSASCSDSGNLGDPLVLSKAPRWHLQRCWCLNFHGRVTVASVKNFOLVAVSDCEA 333
DB 406	QDSDILSRVQKGAGHGLLLQNKAPSWSDSGAYVLNFGHGRVTRASVKNFQIVHPDE--- 462
QY 334	GOTSERIILQFGKVGKDMFTMDYGYPISAFOAFAICLSSFETRIACE 380
DB 463	---PDLVLQGRVAPNIFTMDFRYPLCPLOAFAICLSSFDGKLACE 506
RESULT 2	
US-08-630-592-7	
; Sequence 7, Application US/08630592	
; Patent No. 5770432	
; GENERAL INFORMATION:	
; APPLICANT: Nishina, Patsy	
; APPLICANT: No. 5770432entTrauth, Konrad	
; APPLICANT: Naggett, Juergen	
; APPLICANT: No. 5770432th, Michael	
; TITLE OF INVENTION: Obesity Associated Genes	
; NUMBER OF SEQUENCES: 25	
; CORRESPONDENCE ADDRESS:	
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT	
; STREET: 3400 Embarcadero Center, Suite 3400	
; CITY: San Francisco	
; STATE: California	
; COUNTRY: USA	
; ZIP: 941114187	
; COMPUTER READABLE FORM:	
; MEDIUM TYPE: Floppy disk	
; COMPUTER: IBM PC compatible	
; OPERATING SYSTEM: PCDOS/MSDOS	
; SOFTWARE: Patent In Release #1.0, Version #1.30	
; CURRENT APPLICATION DATA:	
; APPLICATION NUMBER: US/08/630,592	
; FILING DATE:	
; CLASSIFICATION: 435	
; ATTORNEY/AGENT INFORMATION:	
; NAME: Sherwood, Pamela J.	
; REGISTRATION NUMBER: 36,677	
; REFERENCE/DOCKET NUMBER: A59504/BIR/PJS	
; TELECOMMUNICATION INFORMATION:	
; TELEPHONE: (415) 7811989	
; TELEFAX: (415) 3983249	
; TELEX: 910 277299	
; INFORMATION FOR SEQ ID NO: 7:	
; SEQUENCE CHARACTERISTICS:	
; LENGTH: 460 amino acids	
; TYPE: amino acid	
; STRANDEDNESS: double	
; TOPOLOGY: linear	
; MOLECULE TYPE: peptide	
US-08-630-592-7	
Query Match 22.3%; Score 437.5; DB 1; Length 460;	
Best Local Similarity 34.9%; Pred. No. 4.3e-40;	
Matches 109; Conservative 60; Mismatches 82; Indels 61; Gaps 12;	
QY 78	EIVAVPEFSSKLTFFPISLQSGPRDSLVOCFIKRNRNTQ-----SYHLVGLTSLTDN 131

Matches 106; Conservative 44; Mismatches 90; Indels 47; Gaps 7;	
QY 100	PRDSVOCIFKRNNTQ-----SYHLVGLTSLTDNGKFLAASKLRATCTDYILSL 153
DB 261	PRDHVMQWIRVRNKGMDKMPFSYYLYLEGEDGV---AHFLAGRRKRKSTSNYLSL 317
QY 154	RSDDISKRNAVILGRMRSNPLGTFKFTVFDGSGTGAARMQKSRSSNFVKVSPRPQGSYPI 213
DB 318	DPKDSRNGSNFVGKVRSNVLGTFKFTFDNG-----VNPERSYVWPDSAIR-----EL 367
QY 214	AHISYELNVLGRGPRMRCIMDTPMISVSRGVVASTSISFSRSPVFRSHSKPLR 273
DB 368	GVVCYETNVLGRGPRKMTVI-----LPGMDSRKQ---RMKVQPON 405
QY 274	SNSASCSDSGNLGDPLVLSKAPRWHLQRCWCLNFHGRVTVASVKNFOLVAVSDCEA 333
DB 406	QDSDILSRVQKGAGHGLLLQNKAPSWSDSGAYVLNFGHGRVTRASVKNFQIVHPDE--- 462
QY 334	GOTSERIILQFGKVGKDMFTMDYGYPISAFOAFAICLSSFETRIACE 380
DB 463	---PDLVLQGRVAPNIFTMDFRYPLCPLOAFAICLSSFDGKLACE 506
RESULT 2	
US-08-630-592-7	
; Sequence 7, Application US/08630592	
; Patent No. 5770432	
; GENERAL INFORMATION:	
; APPLICANT: Nishina, Patsy	
; APPLICANT: No. 5770432entTrauth, Konrad	
; APPLICANT: Naggett, Juergen	
; APPLICANT: No. 5770432th, Michael	
; TITLE OF INVENTION: Obesity Associated Genes	
; NUMBER OF SEQUENCES: 25	
; CORRESPONDENCE ADDRESS:	
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT	
; STREET: 3400 Embarcadero Center, Suite 3400	
; CITY: San Francisco	
; STATE: California	
; COUNTRY: USA	
; ZIP: 941114187	
; COMPUTER READABLE FORM:	
; MEDIUM TYPE: Floppy disk	
; COMPUTER: IBM PC compatible	
; OPERATING SYSTEM: PCDOS/MSDOS	
; SOFTWARE: Patent In Release #1.0, Version #1.30	
; CURRENT APPLICATION DATA:	
; APPLICATION NUMBER: US/08/630,592	
; FILING DATE:	
; CLASSIFICATION: 435	
; ATTORNEY/AGENT INFORMATION:	
; NAME: Sherwood, Pamela J.	
; REGISTRATION NUMBER: 36,677	
; REFERENCE/DOCKET NUMBER: A59504/BIR/PJS	
; TELECOMMUNICATION INFORMATION:	
; TELEPHONE: (415) 7811989	
; TELEFAX: (415) 3983249	
; TELEX: 910 277299	
; INFORMATION FOR SEQ ID NO: 7:	
; SEQUENCE CHARACTERISTICS:	
; LENGTH: 460 amino acids	
; TYPE: amino acid	
; STRANDEDNESS: double	
; TOPOLOGY: linear	
; MOLECULE TYPE: peptide	
US-08-630-592-7	
Query Match 22.3%; Score 437.5; DB 1; Length 460;	
Best Local Similarity 34.9%; Pred. No. 4.3e-40;	
Matches 109; Conservative 60; Mismatches 82; Indels 61; Gaps 12;	
QY 78	EIVAVPEFSSKLTFFPISLQSGPRDSLVOCFIKRNRNTQ-----SYHLVGLTSLTDN 131

Matches 106; Conservative 44; Mismatches 90; Indels 47; Gaps 7;	
-----------------------------------------------------------------	--

201	EVQDLEFALR-----FAPQGITTKCRITRDKGMDRGMYTFYFLHLD-----RED	246
Qy	132 GK--FLIAASKKRACTDYIISLRSDDISKRSNAYLGRMSRNFILGTFYFDGSGQTGAA	189
Db	247 GKVFLLAGRIRKKSKTSNYLISVDPTDLSRGDSYIGKLSNLMGTKFTYVD-----NGV	302
Qy	190 KMQKSRSNFIKVSFPRVQGSYPYIAHISVELNVLGSRGPRMRCIMDTIP-MSIVESRGV	248
Db	303 NPQKASSTLESGLR-----QELAAVCYETNVLGFKGPRKMSVI-----VPGMNWHER--	352
Qy	249 VASTSISFSRSRSPVFRSHSKPLRSNSASCSDSGNLGDPLVLSNKAPRWHEQLRCWC	308
Db	353 ---VSIRPRNEHETLLARWQNK-----NTESI-----IELQNKTPVWMDTQTSV	394
Qy	309 LNFHGRVTVAASKVQFQLVAVSDCEAGQTSRIILOFGKVGKDMFTMDYGPISAFQAPAI	368
Db	395 LNFHGRVTOASVKVQFQII-----HGNDPDIYVMQFGRVAEDFTMDYNYPLCALQAPAI	448
Qy	369 CLSSFETRIACE	380
Db	449 ALSFFDSKLACE	460

```

RESULT 4
US-09-032-365A-8
; Sequence 8, Application US/09032365A
; Patent No. 6114502
; GENERAL INFORMATION:
; APPLICANT: No. 6114502th, Michael
; APPLICANT: Nishina, Patsy
; APPLICANT: Naggart, Juergen
; APPLICANT: No. 6114502en-Trauth, Konrad
; TITLE OF INVENTION: GENE FAMILY ASSOCIATED WITH
; TITLE OF INVENTION: GENE FAMILY ASSOCIATED WITH
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bozicevic & Reed, LLP
; STREET: 285 Hamilton Avenue, Suite 200
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/032.365A
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: SEQ-2CIP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-327-3400
; TELEFAX: 650 327-3231
; TELEX:
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 460 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-032-365A-8

```

Query Match	22.3%	Score 437.5;	DB 2;	Length 460;
Best Local Similarity	34.9%;	Pred. No. 4.3e-40;		

```

Matches 109; Conservative 60; Mismatches 82; Indels 61; Gaps 12;
Qy 78 EIVAVPEFSSKFLPISLKSGPRDSLQVCFIKRNRNTQ-----SYHLYLGLTTSLTDN 131
Db 201 EVQDLEEFALR-----PAFGITIKCRITRDKKGMDRGMVPTVFLHLD-----RED 246
Qy 132 GK--FLLAASKLRACVCTDYIIISLDDDISKRSNAVILGRMSNFLGTFKFTVFDGSGQTGAA 189
Db 247 GKQVFLLAGRKRKSKTSNYLISVDPDLSRGDSYIGKLSNLMLGTFKFTYD-----NGV 302
Qy 190 KMKKSRSSNFIKVSPPVQGSYPFAHISYELNLVLGSRGPRRMRICIMDTIP--MSIVESRGV 248
Db 303 NPQKASSSTLESGLR-----QELAAVCYETNVLGPKGPKMSVI---VPGMNKMHVR-- 352
Qy 249 VASTSISFSSRSPPVFRSHSKPLRNSASCSDSGNNLGDPPPLVLGNKAPRWHEQLRWC 308
Db 353 ---VSIRPRNEHETLLARQNK---NTESI-----TELQNKTPVWDDTQSYV 394
Qy 309 LNFHGRVTVASVKQFVLAVSDCEAGTSERIILQFGKVKQDMFTWDYGYPISAQAFAI 368
Db 395 LNFHGRVTVASVKQFVLI-----HGNDDPDYIMQFGRVAEDVFTWDYNYPLCALQAFAI 448
Qy 369 CLSSEFETRIACE 380
Db 449 ALSFDSKLACE 460

RESULT 5
US-08-631-200-8
; Sequence 8, Application US/08631200
; Patent No. 5646040
; GENERAL INFORMATION:
; APPLICANT: Kleyn, Patrick W.
; APPLICANT: Moore, Karen J.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/631,200
; FILING DATE: 12-APR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-057
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8664
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 506 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-631-200-8

```

Query Match 22.3%; Score 437.5; DB 1; Length 506;
Best Local Similarity 34.9%; Pred. No. 5e-40;
Matches 109; Conservative 60; Mismatches 82; Indels 61

[illegible]

THIS PAGE BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 23, 2005, 23:24:03 ; Search time 163 Seconds
(without alignments)
974.081 Million cell updates/sec

Title: US-10-763-042-9

Perfect score: 1962

Sequence: 1 MFRSLQEMRSPRHVVHA.....SAFQFAICLSFETRIACE 380

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep.*
- 3: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep.*
- 4: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 5: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 6: /cgn2_6/prodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1962	100.0	380	5	US-10-763-042-9
2	1414.5	72.1	380	5	US-10-763-042-11
3	1149	58.6	406	5	US-10-763-042-3
4	1132	57.7	391	4	US-10-424-599-149493
5	1103	56.2	400	4	US-10-424-599-284656
6	1035.5	52.8	407	4	US-10-437-963-193011
7	1023.5	52.2	445	4	US-10-437-963-160306
8	1018	51.9	436	4	US-10-425-114-59800
9	1018	51.9	436	4	US-10-425-114-72469
10	1016	51.8	448	5	US-10-732-923-19498
11	1013.5	51.7	455	5	US-10-732-923-19480
12	1011	51.5	450	4	US-10-767-701-41195
13	1010.5	51.5	437	4	US-10-425-114-57097
14	1010.5	51.5	437	4	US-10-425-115-202431
15	1009.5	51.5	445	4	US-10-437-963-150007
16	1006	51.3	442	4	US-10-425-114-59799
17	1005.5	51.2	631	5	US-10-732-923-19497
18	999.5	50.9	445	4	US-10-225-068-130
19	999.5	50.9	445	4	US-10-374-780A-2896
20	999.5	50.9	445	5	US-10-763-042-10
21	999.5	50.9	445	5	US-10-732-923-19479
22	999.5	50.9	445	5	US-10-225-068-130
23	992	50.6	448	4	US-10-437-963-199678
24	992	50.6	448	5	US-10-732-923-19494
25	991.5	50.5	450	5	US-10-763-042-1
26	989	50.4	440	5	US-10-732-923-19495
27	989	50.4	440	5	US-10-732-923-19496

28	989	50.4	462	4	US-10-437-963-153289	Sequence 153289, A
29	989	50.4	468	4	US-10-425-114-62614	Sequence 62614, A
30	985	50.2	494	4	US-10-425-115-355201	Sequence 355201, A
31	972	49.5	432	5	US-10-732-923-19488	Sequence 19488, A
32	972	49.5	527	4	US-10-425-114-59801	Sequence 59801, A
33	971	49.5	440	4	US-10-425-115-307562	Sequence 307562, A
34	971	49.5	551	4	US-10-425-114-65527	Sequence 65527, A
35	969	49.4	440	4	US-10-437-963-128060	Sequence 128060, A
36	966.5	49.3	394	5	US-10-763-042-2	Sequence 2, Appli
37	966.5	49.3	427	4	US-10-424-599-151222	Sequence 151222, A
38	966.5	49.3	429	5	US-10-732-923-19487	Sequence 19487, A
39	966	49.2	440	5	US-10-732-923-19499	Sequence 19499, A
40	955	48.7	480	4	US-10-310-154-658	Sequence 658, App
41	955	48.7	480	5	US-10-732-923-612	Sequence 612, App
42	950.5	48.4	429	5	US-10-763-042-5	Sequence 5, Appli
43	950.5	48.4	429	5	US-10-732-923-19484	Sequence 19484, A
44	949	48.4	414	4	US-10-424-599-254939	Sequence 254939, A
45	937	47.8	428	5	US-10-732-923-19510	Sequence 19510, A

ALIGNMENTS

RESULT 1

US-10-763-042-9
; Sequence 9, Application US/10763042
; Publication No. US20050014266A1
; GENERAL INFORMATION:
; APPLICANT: Shaw, Jei-Fu
; APPLICANT: Lai, Chia-Ping
; TITLE OF INVENTION: PLANT TUBBY-LIKE PROTEINS
; FILE REFERENCE: 08919-099001
; CURRENT APPLICATION NUMBER: US/10/763,042
; CURRENT FILING DATE: 2004-01-21
; PRIOR APPLICATION NUMBER: US 60/441,380
; PRIOR FILING DATE: 2004-01-21
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Arabidopsis sp.
US-10-763-042-9

Query Match	100.0%;	Score 1962;	DB 5;	Length 380;
Best Local Similarity	100.0%;	Pred. No. 5.6e-192;		
Matches 380;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MFRSLQEMRSPRHVVHA	A	A
Db	1	MFRSLQEMRSPRHVVHA	A	A
Qy	61	NNVACAGVCRWRILTK	KEIVAVPEFSSKLT	FFPISLQSGPRDSLVC
Db	61	NNVACAGVCRWRILTK	KEIVAVPEFSSKLT	FFPISLQSGPRDSLVC
Qy	121	YLGTLTSLDNGKFLA	ASKLRATCTDVI	ISLRSDDISKRSNAYL
Db	121	YLGTLTSLDNGKFLA	ASKLRATCTDVI	ISLRSDDISKRSNAYL
Qy	181	FDGSQTGAAMQKRS	SNFIKUSPRVPGQSY	PIAHISYELNVLGSR
Db	181	FDGSQTGAAMQKRS	SNFIKUSPRVPGQSY	PIAHISYELNVLGSR
Qy	241	SIVESRGVAVSTISS	FSRSRSPVFRSHKPL	RSNSASCDSGNNLGD
Db	241	SIVESRGVAVSTISS	FSRSRSPVFRSHKPL	RSNSASCDSGNNLGD
Qy	301	HEQLRCWCLNFHGR	VTVASVQNFQLVAV	SDCEAGQTSERIILO
Db	301	HEQLRCWCLNFHGR	VTVASVQNFQLVAV	SDCEAGQTSERIILO
Qy	361	SAFQFAICLSFET	RIACE 380	

Qy	59	RNVVACAGVCSWRILTKBIIVAVPEFSGSKITFPISLKGSGPRDSLVOCTIKNRNTQSY	118
Db	62	RKWSVACAGVCSWRHITKDIIVKTPELSGSKITFPISVKGQPRENLARCIPKRNSTQTY	121
Qy	119	HLVGLHTTSLDNGKFLLAASKLKRATCTDVIISLRSDDISKRNAYLGRMRNFIJGTFK	178
Db	122	YLFSLSTSLAEDGKFLLAARKCRPTCTDVIISLDADMSKGSNVVGLKRNFIJGTFK	181
Qy	179	TVPDGS--OTGAANKQKGRSS---NFIKVSPRVPOG-SYPIAHISYELNV-LGSRGPRM	231
Db	182	TIYDSLQHPHG-AKTMKSRSTRLNVLKQVSPKVTGXNYPVAHISYEFKXKCLGSRGPRM	240
Qy	232	RCIMDTIPMSIVBSRGVASTSISFSFSS-----RSPVFRSHSKPLRNSASCDSGN--N	285
Db	241	HCVMDSIPATAIEPGGVAPTQFTEFLSNIDMFPSPFPFPRS-----KSNRAENSMGSLVD	295
Qy	286	LGDPPPLVLANKAPRWHEQLRCWCLNFHGRVTVASVXNQLVA-VSDCEAGTSERIILQF	344
Db	296	QKDGMLVLKNKAPRWHEQLRCWCLNFHGRVTIASVKNFQLVASAENGCPAGPEHDKIILQF	355
Qy	345	GKVGKDMFTMDYGYPISAQPAFACLSSFETRIACE	380
Db	356	GKVGKOLFTMDYRYPISAFQPAFACLSSFDFTRKIACE	391

RESULT 5

```

US-10-424-599-284656
; Sequence 284656, Application US/10424599
; Publication NO. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 284656
; LENGTH: 400
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_9906C.1.pep
; US-10-424-599-284656

```

RESULT 6

```

US-10-437-963-193011
; Sequence 193011, Application US/10437963
; Publication No. US20040123343A1
;
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
;
; SEQ ID NO 193011
; LENGTH: 407
; TYPE: PRT
; ORGANISM: Oryza sativa
;
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(407)
; OTHER INFORMATION: unsure at all Xaa locations
;
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_8918C.1.pep
;
US-10-437-963-193011

```



```
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3591-079-H7_FLI pep
US-10-425-114-72469

Query Match      51.9%; Score 1018; DB 4; Length 436;
Best Local Similarity 48.3%; Pred. No. 4.5e-95;
Matches 213; Conservative 65; Mismatches 97; Indels 66; Gaps 10;

QY 1 MFRSLLOEM-----RSRPHRVVHAAASTANSSDPFSSSELPEEL 40
Db 1 MSFRSIVRDVDFGSLSRGFEVKFLGHRGRSGHGAHVELRDPAPVIOSSCWANLPPEL 60

QY 41 LREILIRVTDGDDWPSRNVVACVCSWRILTKETIVAVPEFSKLTFFPISLKOSGP 100
Db 61 LRDVIERLEASE-ATWPNRKSVVSCAAVCKTWREICKEIVKNPEFSKLTFFPISLKOSGP 119

QY 101 RDSLVQCFIKRNRNTQSYHLYGLTTS-LTDNGKFLLAASKLRATCTDYIISLRDDIS 159
Db 120 RDVTIOCFIKRDKSTQTYIYLYLCLSTAVLVEGKFLCAKTSRPTCTEYTFPMSENLS 179

QY 160 KRSNAYLGRMSNFLTGTFTVDFGSGTGAAMQKRSNFI--KVSPPRVQGSYPYIAHIS 217
Db 180 RSKMVIKLSNLLGTKEIAYDTQPPCNAABPKTSRRFYKRVKSPKVSSTYNIAQVS 239

QY 218 YELNVLGSRGPRMRCIMDTIPMSIVESRGVVAS-----TSISFSS-SRSSPVP 265
Db 240 YELNVLGTPRRMNCVMSIPTSCLAEAGTVPSQLDSILARSTDSFGSISFSKSVVD 299

QY 266 RSHSKPLRSNSASCDP-----SGNNLQDP-----PLVLSNKAAPRWHEQLRCWCLN 310
Db 300 RS-----WRSFSSRFSDISSSRRIGDTASGDNDECKERPLILRNKAAPRWHEQLRCWCLN 355

QY 311 FHGRVTVASVKNQFQVAVSDCEAGQTS-----BRILQFGKVGKMDTMDYGYTP 359
Db 356 FGRVTVASVKNQFQVAVSDCEAGQTS-----BRILQFGKVGKMDTMDYGYTP 415

QY 360 ISAPQAFALCLSSPFTRIACE 380
Db 416 LSAFOAFAICLSSFDTKLACE 436

RESULT 10
US-10-732-923-19498
; Sequence 19498, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 19498
; TYPE: PRT
; LENGTH: 448
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(448)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-732-923-19498

Query Match      51.8%; Score 1016; DB 5; Length 448;
Best Local Similarity 49.9%; Pred. No. 7.5e-95;
Matches 206; Conservative 63; Mismatches 100; Indels 44; Gaps 8;

QY 11 RSRPHRVVHAAASTANSSDPFSSSELPEELRILIRVTDGDDWPSRNVVACVCK 70
Db 37 RGSQSTVHELCDADLIIOESRWASLPPELRLDVIIRLSEASTXWPSRNDVVSCAACK 96

QY 71 SWRILTKETIVAVPEFSKLTFFPISLKOSGPRDLSVOCFIKRNPTOSYHLYGLTTS-LT 129
Db 97 AWRECKEIVLSPFCGKLTFFPISLKQPGPRDGMIOCFIKRDKSKSYHLYLCLSTAVLA 156

QY 130 DNGKFLLAASKLRATCTDYIISLRDDISKRSNAYLGRMSNFLTGTFTVFD-----G 183
Db 157 DSGKFLLSAKRHKTKTCTEYVISMADNISRSGSTYIGKLSNFLTGTFTKIYDTQPSYNG 216

QY 184 SQTGAAMQKRSNFIKVSPPRVQGSYPYIAHISYELNVLGSRGPRMRCIMDTIPMSIV 243
Db 217 AVIPPVXRSRRSFRNSKVSPPKMPGSYINIAQVYELNVLGTRGPRMRCVMSIIPASSV 276

QY 244 ESRGVV-----ASTSISFSSRSPPVRS---HSPKPLRSNSASCSDSGNNL 286
Db 277 EFGIVPGQPEQIVPAPFESFRSTTSP-SKSIDRSMGFOXPDXDFSARSFSDIAGGT 335

QY 287 GDP-----PLVLSNKAAPRW-HEQLRCWCLNFGHRTVTVASVKNQFQVAVSDCE----- 332
Db 336 SDBEGQNKERPLVLRNKAAPRWHEQLQCWCLNFRGRVTIASVKNQFQVAVSDCE----- 395

QY 333 -----AGTSEIRILLOFGKVGKMDTMDYGYPIASQAFALCLSSPFTRIACE 380
Db 396 TPSQPAPPQDKIILQFGKVAKMDTMDYRPLSAFOAFAICLSSFDTKLACE 448

RESULT 11
US-10-732-923-19480
; Sequence 19480, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 19480
; TYPE: PRT
; LENGTH: 455
; ORGANISM: Arabidopsis thaliana
US-10-732-923-19480

Query Match      51.7%; Score 1013.5; DB 5; Length 455;
Best Local Similarity 47.1%; Pred. No. 1.4e-94;
Matches 215; Conservative 64; Mismatches 100; Indels 77; Gaps 10;

QY 1 MFRSILLQEMR---SRPHRVVHAAASTANSSDPFS-----WSEL 36
Db 1 MSFRSIVRDVDSIGLSRRSDFKLSSLNKGKSGSVQDSHESHEQLVVTIQTETPANL 60

QY 37 PEELLREILIRVTDGDDWPSRNVVACVCSWRILTKETIVAVPEFSKLTFFPISLK 96
Db 61 PEELLRDVTKLEESE-SVWPARRHVACVCSWRDMCKEIVQSPESLTKITFFVSLK 119

QY 97 QSGPRDLSVQCFIKRNRNTQSYHLYGLTTS-LTDNGKFLLAASKLRATCTDYIISLRS 155
Db 120 QGPRDATWQCFIKRDKSNLTVHLYLCLSPALLVENGKFLLSAKIRRTTYEYVISMHA 179

QY 156 DDISKRSNAYLGRMSNFLTGTFTVFD-----GSQTGAAMQKRSNFI--KVSPPRVQ 208
Db 180 DTISRSNTYIGKIRSNFLTGTFTYDTOPAYNSIARAVQPVGLSRRFYKRVSPKVPFS 239

QY 209 GSYPIAHISYELNVLGSRGPRMRCIMDTIPMSIVESRGV-----ASTS 253
Db 240 GSYKIAQVSYELNVLGTRGPRMRCVMSIPTSCLAEAGTVPGQPDIIIVPSRILDSFSS 299

QY 254 ISSFSR-----SSPVSFRSHSKPLRSNSASCSDSGNNLQDPPLVLSNKAAPRWHEQ 303
Db 300 ITSSSRKITYDYSNDFSSARSFSDILGPLSEDEQEVVLESGKERNSPVLKKNKPPRWHEQ 359

QY 304 LRCWCLNFGHRTVTVASVKNQFQVAVSDCE-----AGQTS---ERILQF 344
```


US-10-425-115-202431

Query Match 51.5%; Score 1010.5; DB 4; Length 437;
Best Local Similarity 48.0%; Pred. No. 2.6e-94;
Matches 212; Conservative 66; Mismatches 97; Indels 67; Gaps 10;
QY 1 MTFRSLQRM-----RSRPHRVHAAASTANSSDPFWSLPEEL 40
DB 1 MFRSIVRVRDGFGLSRGRFEVFLGHRGRSHGAVHELDPAPVIOSSCWANLPPEL 60
QY 41 LREILIRVTVGGDWPSRRNVVACAGVCRSWRIITKEIVAVPEFSSKLTFFPISLKQSGP 100
DB 61 LRDVIERLEASB-ATWPNKSVVSCAAVCRTWREICKEIVKNPEFSGKITFPVSLKQCP 119
QY 101 RSLVQCFIKRNRNTQSYHYLGLTTS-LTDNGKFLLAASKUKRATCTDYIISLRSDDIS 159
DB 120 RDVTIQCFIKRDKSTQTYLYLCLSTAVLVESGKFLCAKTSRPTCTEYTFPMNSNL 179
QY 160 KESNAYLGRMRNFLTGTFTVDSQGTGAQKQKSRSSNFI--KVSPPVQSGSYPTAHTS 217
DB 180 RSSKMVIGKRLNLTGTFKFAIYDTQPPCNAAPGKTSRRFYRKVSPKVSSTYNTAQS 239
QY 218 YELNVLGSRPRMRMRCIMDTIPMSIVESRGVVAS-----TSISFS-SRSSPVF 265
DB 240 HELNVLGTGPRRMVNCVMSHSIPTSCLEAGSVPSQLDSILARSTDESFGSISFSKSSVD 299
QY 266 RSHSKPLRSNSASCSD-----SGNNLGD-----PLVLSNKAPRWHEQLRCWCIN 310
DB 300 RS-----MRFSRFRSDISMSRRIGDTASGDNDECKERFLILRNKAPRWHEQLRCWCIN 355
QY 311 FGRVTVASVKNFOLVAVSDCEAGQTS-----ERILQFGVKGMFTMDYGY 358
DB 356 FGRVTVASVKNFOLVAAOQPAAGAPTPSOATVPVPEHEKVILOFGKVAKDMFTMDYRY 415
QY 359 PISAFQAFALCLSSPETRIACE 380
DB 416 PLSAFQAFALCLSSFDTKLACE 437

RESULT 15

US-10-437-963-150007
; Sequence 150007, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 150007
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_50284C.1.pgp
US-10-437-963-150007

Query Match 51.5%; Score 1009.5; DB 4; Length 445;
Best Local Similarity 51.1%; Pred. No. 3.4e-94;
Matches 210; Conservative 66; Mismatches 92; Indels 43; Gaps 11;
QY 11 RSRPHRVHAAASTANSSDPFWSLPEELIRIVETVDGDPWPSRRNVVACAGYCR 70
DB 37 RGKSQSTVHELCDADLIIOESRWASLPPELLRDVIRRLASE-STWPSRKDVVSCAAVCK 95

QY 71 SWRIITKEIVAVPEFSSKLTFFPISLKQSGPRDSLVCQFIKRNNTQSYHYLGLTTS-IT 129
DB 96 AWRECKEIVLSPFCGKLTFFPLSLKQPGPRDGMICQFIKRDKSKSTYHYLCLSTAVLA 155
QY 130 DNGKFLLAASKUKRATCTDYIISLRSDDISKSNAYLGRMRNFLTGTFTVDP--GSQTG 187
DB 156 DSGKFLLSAKRHKTKTEYVISMADNISRSSSYIGKLRNFLTGTFTIYDTQPSYNG 215
QY 188 AAKMQKSRSS---NFIVKSPRVPOQSGSYPTAHTSYELNVLGSRGPRMRMRCIMDTIPMSIVE 244
DB 216 AVIPPVGRSSRRFNSKVSPPKSPGSYNTAQTVELNVLGTRGPRRMHCVMSIPASSVE 275
QY 245 SRGVV-----ASTSISFSRSPVFRS--HSKPLRSNSASCSD-----S 282
DB 276 PGGIIVPGQPEQIVPRAFEESFRSTTSF-SKSSIMDRSMDFFSSRDFSSARFSDIAGGTIN 334
QY 283 GNNLG---DPPLVLSNKAPRWHEQLRCWCINLPHGRVTVASVKNFOLVAV-SDCEAGQTS- 337
DB 335 GDEEGONKERPLVLRNKAPRWHEQLRCWCINLPHGRVTVASVKNFOLVAV-SDCEAGQTS- 394
QY 338 -----ERILQFGVKGMFTMDYGYPIISAFQAFALCLSSPETRIACE 380
DB 395 SQPAPPEODKIILQFGKVAKDMFTMDYRYPLSTFQAFALCLSSFDTKLACE 445

Search completed: December 23, 2005, 23:32:33
Job time : 165 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 23, 2005, 23:12:06 ; Search time 39 Seconds
(without alignments)
937.497 Million cell updates/sec

Title: US-10-763-042-9
Perfect score: 1962
Sequence: 1 MFRSLQEMRSPRHVVHA.....SAFOAFLCSSFETRIACE 380

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80.*
1: PIR1.*
2: PIR2.*
3: PIR3.*
4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1147.5	58.5	407	2 H84920	probable Tub famil
2	1013.5	51.7	455	2 H96797	hypothetical prote
3	999.5	50.9	445	2 E86382	hypothetical prote
4	929	47.3	388	2 E96513	unknown protein, 3
5	919.5	46.9	386	2 E84562	probable Tub famil
6	787.5	40.1	415	2 F96499	hypothetical prote
7	437.5	22.3	271	2 S42728	phosphodiesterase
8	436.5	22.2	505	2 S68518	tub protein, brain
9	394.5	20.1	425	2 T20691	hypothetical prote
10	352.5	18.0	265	2 T02138	hypothetical prote
11	201	10.2	397	2 E86295	hypothetical prote
12	109.5	5.6	360	2 H86347	hypothetical prote
13	106.5	5.4	695	2 D84634	hypothetical prote
14	103	5.2	322	2 H85068	N7-like protein [i
15	98.5	5.0	391	2 H85063	hypothetical prote
16	98.5	5.0	775	2 T49957	hypothetical prote
17	98.5	5.0	1360	2 T12064	DNA binding protei
18	96	4.9	1161	2 S57180	probable membrane
19	94.5	4.8	1812	2 T49350	breast/ovarian can
20	94	4.8	1532	2 H96795	hypothetical prote
21	93.5	4.8	383	2 F96671	hypothetical prote
22	93.5	4.8	1164	2 S71792	phosphatidylinosi
23	92.5	4.7	736	2 S45859	hypothetical prote
24	92	4.7	542	2 S64030	probable membrane
25	92	4.7	812	2 T52569	squamosa-promoter
26	92	4.7	978	2 H86319	hypothetical prote
27	92	4.7	1035	2 G86342	hypothetical prote
28	91.5	4.7	1295	2 G96549	hypothetical prote
29	90.5	4.6	567	2 JC5957	transforming growt

30	90.5	4.6	506	2 JC5956	transforming growt
31	90	4.6	560	2 T40608	conserved hypotet
32	89.5	4.6	1160	2 T23713	hypothetical prote
33	89.5	4.6	2144	2 S71490	ashi protein - fru
34	89	4.5	1075	2 S48992	floculation prote
35	88.5	4.5	1206	2 E86445	hypothetical prote
36	88	4.5	1131	2 T41144	hypothetical serin
37	87.5	4.5	1261	2 T50065	hypothetical prote
38	87.5	4.5	2833	2 A43360	inositol 1,4,5-tri
39	87	4.4	418	2 D86241	protein T16B5.8 [i
40	87	4.4	1286	2 T33476	hypothetical prote
41	87	4.4	2609	2 T40399	probable transport
42	86.5	4.4	393	2 AC2892	acyl-CoA dehydrog
43	86.5	4.4	393	2 F97667	probable carnitine
44	86.5	4.4	579	2 JC5955	transforming growt
45	86.5	4.4	779	2 S64680	exoribonuclease 10

ALIGNMENTS

RESULT 1

H84920
probable Tub family protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 31-Dec-2004
C:Accession: H84920
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; N
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: H84920
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-407 <STO>
A:Cross-references: UNIPROT:082257; UNIPARC:UPI00000A8946; GB:AE002093; NID:g3738302; PII
C:Genetics:
A:Map position: 2
A:Gene: At2g47900
C:Superfamily: F-box protein tubby-like, plant type

Query Match	58.5%	Score	1147.5;	DB	2;	Length	407;
Best Local Similarity	56.6%;	Pred.	No. 5.6e-92;				
Matches	233;	Conservative	57;	Mismatches	85;	Indels	37;
Gaps	9;						
QY	1	MFRSLQEMRSPRHVVHA	-----RSRPHRVVHAAASTANSSDPFNSLPEELL	41			
Db	1	MSFKSLIQMRGELGISIRKGFDFRGYSRQVVQDTSPVDFAFKSCWASMPPELL	60				
QY	42	REILIRVETVDCGDWPSRRNVACAGVCRSWRLTKEIVAVPEFFSKLTFPISLKQSGPR	101				
Db	61	RDVLMRIEQSE-DTWPSRKNNVSCAGVCNREIVKEIVRVELLSKLTFFISLKQSGPR	119				
QY	102	DSLVCQFKIRNRTQSYHLYLGL--TTSLTUNGKFLAASKLKRATCTDYIISLRDDIS	159				
Db	120	GSLVCQYIMRNRNQTYLYLGLNQAAASNDGKFLAAKRRFRPTCTDYIISLNCDDVS	179				
QY	160	KESNAYLGMRNFIKGTFTVEDGSGQTG-AAKMQKSRSSNFI---KVSRRVPGSGYPIAH	215				
Db	180	RSNTYIGKLRNFIKGTFTVYDAQPTNPGTQVTRSRRLSLKQVSPRIPSGNTPVVAH	239				
QY	216	ISYELNVLGSRGPRMRRCIMDTIPMSIVBSRGV-----VASTSISSFSRSPSPVFRSHS	269				
Db	240	ISYELNVLGSRGPRMRRCIMDTIPMSIVBSRGV-----VASTSISSFSRSPSPVFRSHS	295				
QY	270	KPLRNSASCSGDSGNLGDPLVLNKNKAPRWHEQLRCWCNLFHGRVTVVASVKNFQIVAV-	328				
Db	296	KGIRAESLPSGSSAAQKGLLVNKNKAPRWHEQLRCWCNLFHGRVTVVASVKNFQIVAAAP	355				
QY	329	SCEAGQTSERILIQGVKGMFTMDYGYPIASAQAFALCISSFFETRIACE	380				

A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: E96513
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-388 <STO>
A:Cross-references: UNIPROT:Q9C6B4; UNIPARC:UPI00000A18D9; GB:AE005173; NID:g10092408; P
C:Genetics:
A:Map position: 1
C:Superfamily: F-box protein tubby-like, plant type

Query Match 47.3%; Score 929; DB 2; Length 388;
Best Local Similarity 50.8%; Pred. No. 5.5e-73;
Matches 197; Conservative 65; Mismatches 92; Indels 34; Gaps 12;

QY 11 RSRPH-----RVHAAASTANSDDPSWSELPELLEIREILIRVETVDDGDMPSRRNVVAC 65
DB 17 RGRSHIAPEGSSVSSLSLSTNEGLNQSIWDLPELLELDIIQRIES-EQSLWPGRRDWWAC 75
QY 66 AGVCRSMRILTKETVAVPSSKLTTPISIKQSGPRDSLVOCEIKENRNTQSYHLVGLT 125
DB 76 ASVCKSWRETKVVKVPELSGLITTPISLROQPRDAPICQPKRERATGIYRLVGLS 135
QY 126 TSIT-DNGKFLAASKRATCTDYIISLRSDDISKRSNAYLGRMSNPLGTFTVFDGS 184
DB 136 PALSGDKSKLLSAKVRATGAEFVVVLSGNDFSRSSNYIGKLSNPLGTFTVYENQ 195
QY 185 QTG-AAAMQKSRSNFIKVSFPRV--PQSGYPPIAHISYELNVLSGRRPRMRCIMDTIPMS 241
DB 196 PPPFNRLKPPS-----MQVSPWYSSSSSYNTASILYELNVLRTRGPRMRCIMHSIPIS 250
QY 242 IVESRGVASTISISSPSSSSPVFRSHSKPLSNSASCSDS--GNLGDPPPLVLSNKA 299
DB 251 AIOEGGKIQSP--TEFTNQ-----KKKKPLMD---FCSGNLGGESVKEPILILKNKSPR 301
QY 300 WHEQLRCWCLNFHGRVTVASVKNFQVAVSDCEAG-----QTSERIILQFGKVGKDMF 352
DB 302 WHEQLQCWCLNFGRVTVASVKNFQVAVAA--AAGKGNMIPBEEQDRVILQFGKIGDIP 360
QY 353 TWDYGYPISAFAQFAICLSSFFETRIACE 380
DB 361 TWDYRYPISAFQFAICLSSFFETKPACE 388

RESULT 5
E94562
probable Tub family protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 31-Dec-2004
C:Accession: E94562
R:Lin, X.; Kaul, S.; Rounleley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: E94562
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-386 <STO>
A:Cross-references: UNIPROT:Q9ZPW1; UNIPARC:UPI00000A83F0; GB:AE002093; NID:g4309738; P
C:Genetics:
A:Map position: 2
C:Superfamily: F-box protein tubby-like, plant type

Query Match 46.9%; Score 919.5; DB 2; Length 386;
Best Local Similarity 49.4%; Pred. No. 3.7e-72;
Matches 198; Conservative 59; Mismatches 107; Indels 37; Gaps 8;

QY 1 MTFRSILQEMR-----SRPHRVHAAASTANSDDPS--WSELPELLEIREIL 45
DB 1 MSLSKILRLKVEVDGLGISKRSWSKSHIAPDQTPPLDNIPOSPWASLPPELLHDI 60

QY 46 IRVETVDDGDMPSRRNVVACAGVCRSMRILTKETVAVPSSKLTTPISIKQSGPRDSL 105
DB 61 WRVSEETA-WPAAAVSCASVCKSWRGITMEIVRIPEQCGKLTTPISLUKQGGPRDSPI 119
QY 106 QCFIKENRNTQSYHLVGLTTSILNTONGKFLAASKLKRATCTDYIISLRSDDISKRSNAY 165
DB 120 QCFIKENRATATYILYGLMPSETENDKLLAARRIRRTACTDFIISL-----SSLL 171
QY 166 LGRMSNPLGTFTVDPDG--SQTGAAMQKSRSNFIKVSFPRVQSGYPPIAHISYELNVLG 224
DB 172 ILWDRSGFLGTFTTYDNQTASTAQAPNRLHPKQAAPKLPTNSSTVGNITVELNVLR 231
QY 225 SRGPRMRCIMDTIPMSIVESRGVASTISISSPSSSSPVFR--SHSKPLSNSASCSDS 282
DB 232 TRGPRMRCIMDTIPMSIVESRGVASTISISSPSSSSPVFR--SHSKPLSNSASCSDS 287
QY 283 GNNLGDPPPLVLSNKAAPRHEQLRCWCLNFHGRVTVASVKNFQVAVSDCEAG---QTSER 339
DB 288 ---LRDQPLVKNKSPRHEQLQCWCLNFGRVTVASVKNFQVAVSDCEAG---QTSER 344
QY 340 IILQFGKVGKDMFTWDYGYPISAFAQFAICLSSFFETRIACE 380
DB 345 VILQFGKIGKIDFTWDYRYPISAFQFAICLSSFFETKPACE 385

RESULT 6
F96499
hypothetical protein T10P12.9 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2004
C:Accession: F96499
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.P.; Hughes, B.; Huiziar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: F96499
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-415 <STO>
A:Cross-references: UNIPROT:Q9XIF9; UNIPARC:UPI00000A5392; GB:AE005173; NID:g5080765; P
C:Genetics:
A:Map position: 1
C:Superfamily: F-box protein tubby-like, plant type

Query Match 40.1%; Score 787.5; DB 2; Length 415;
Best Local Similarity 43.2%; Pred. No. 1.3e-60;
Matches 168; Conservative 65; Mismatches 87; Indels 69; Gaps 9;

QY 33 WSELPELLEIREILIRVETVDDGDMPSRRNVVACAGVCRSW---RIILTKEIVAVPEFSSKL 89
DB 55 WANLPAALLRDVWKLDSE--STWPA-----RSMFRYLLPFDIMSV-----I 96
QY 90 TTPISLUKQSGPRDSLVOCEIKENRNTQSYHLVGLTTS--LTONGKFLAASKLKRATCTD 148
DB 97 LTTLLMVKPGPRDGIQCYIKRDKSNMTYHLVLSLSPAILVSGKFLLSAKRSRRATYTE 156
QY 149 YIISLRSDDISKRSNAYLGRMSNPLGTFTVFD-----GSQTGAAMQKSRSNFIKVS 203
DB 157 YVISMADNISRSSSYIGKLSNPLGTFTFYDTAPAYNSQILGPPNRSRFSNKKVS 216
QY 204 PRVQSGYPPIAHISYELNVLSGRRPRMRCIMDTIPMSIVESRGVASTISISSPSSSSP 263
DB 217 PKVPSGSYNIAQVTEYELNLLGTRGPRMRCIMHSIPSLALEPGGTYP-----SQP 266

Db 231 AKVRSNALGQTFTVDSGQ-----NPKKTNHAIROE-----LAAVIYETNVLGFK 277
 Qy 227 GPRMRCIMDTIPMSIVESRGVASTSISFSRSPVFRSHGKPLRSNASCSDSGNL 286
 Db 278 GPRKWTIVMP-----GIBPPT-----ENRPAVRCVPRPIQDKHTLLERYLND 320
 Qy 287 GDPPLVLSNKAPRWHEQLRCWCLNFHGRVTVASVKNFOLVAVSDCEAGQTSERIILOFGK 346
 Db 321 LDSLKILSNKSPQWNETQSVLNFHGRVTVQASVKNFQIIHQ-----SPEYIVMQFGR 374
 Qy 347 VGKDMFTMDYGPISAFQAFALCISSPETRIA 378
 Db 375 ISDDEFTMDFRYPLSAVQAFGIAMTSPHGKLA 406

RESULT 10
 T02138
 hypothetical protein F8K4.13 - Arabidopsis thaliana
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 09-Jul-2004
 C/Accession: T02138
 R/Vysotskaya, V.S.; Schwartz, J.R.; Toriumi, M.; Kwan, A.; Yu, G.; Oji, O.; Liu, S.; Li, A.;
 rtz, D.; Li, Y.; Palm, C.J.; Shinn, P.; Sun, H.; Davis, R.W.; Ecker, J.R.; Federspiel, N.
 submitted to the EMBL Data Library, August 1998
 A/Description: Arabidopsis thaliana chromosome 1 BAC F8K4 sequence.
 A/Reference number: Z14574
 A/Accession: T02138
 A/Status: translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-265 <VYS>
 A/Cross-references: UNIPROT:O80699; UNIPARC:UP100000AAA02A; EMBL:AC004392; NID:G3282170;
 C/Genetics:
 A/Name: ATSP:F8K4.13
 A/Map position: 1
 A/Introns: 61/3; 110/1; 199/3

Query Match 18.0%; Score 352.5; DB 2; Length 265;
 Best Local Similarity 29.3%; Pred. No. 5e-23;
 Matches 105; Conservative 36; Mismatches 76; Indels 139; Gaps 12;

Qy 36 LPBELLREILIRVETVDGGPBRNVVACAGVCRSHRLTKEIVAVPESSKLTTPISL 95
 Db 1 MPPELLKDLVLRISERSE-DTWPRKKNVSVCGVKCKMWRQIFKEIVNVPVSSKFTPPISL 59
 Qy 96 KQSGPRDSLVCQCFKRNRTQSHVLYGLTSLTDNGKFLLAASKLRATCTDYIISLRS 155
 Db 60 KQPGGSLVQCYVKNRSNQTFLYL----- 87
 Qy 156 DDISKRSNAYLGRNRSNFLTGTFTVDPGSGTAAMQKRSNFIKVPVPGSGVPIAH 215
 Db 88 -----GEAKI-----FCQSEPSEPKNK----- 104
 Qy 216 ISVELNVLGSGPRMRCIMDTIPMSIVESRGVVASTSISPSRSPVFRSHGKPLRSN 275
 Db 105 -----IWKLKLPGGTATTQTELDNFVSPRSP----- 131
 Qy 276 SASCSDSGNLGDPLVLSNKAPRWHEQLRCWCLNFHG-RVTVASVKNFOLVAVSDCEAG 334
 Db 132 -----SQKEG--VLVLKSKVPLEEQ--SWCLDFNGWRDVISSGKKFQLVALL----- 176
 Qy 335 QTSERIILQFGVKGMFTMDYGY-----PISAFQAF-----IC-----LSSPETRIA 378
 Db 177 RTNLRMKTTFSSLRKSETCTNSSYEAWIPLVRTSVFAVIARVCRDKKHTPSYELKLA 234

RESULT 11
 E86295
 hypothetical protein T24D18.17 - Arabidopsis thaliana
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2004
 C/Accession: E86295
 R/Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K. A.;

ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.; Li, J.H.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, R.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: B86295
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-397 <STO>
A:Cross-references: UNIPROT:Q9S9M8; UNIPARC:UPI00000AAAF61; GB:AE005172; NID:g6587813; PII
C:Genetics:
A:Map position: 1
C:Superfamily: tubby

Query Match 10.2%; Score 201; DB 2; Length 397;
Best Local Similarity 24.4%; Pred. NO. 1.4e-09;
Matches 87; Conservative 54; Mismatches 103; Indels 112; Gaps 18;

Qy 70 RSWRLATKE---IVAVPEFSKLTFFPISLKQSGPRDSLIV-----QCFIKENRNT 115
Db 106 KTWTSVDSEHSSLUKWEFSDSEAPASSWSLTNRLCKTLPDVGRCCTCLIVKEQSP 165
Qy 116 QS-----YHLY-----LGLTSLTDNGKFL--AASKLKRATCTDYIISLRSDD 157
Db 166 EGLSHGSVYSILYHEGRCKDKLAVAHSRNGKSIFFVAQNVKGLLCS-----SDE 218
Qy 158 ISKRSNAYLGRMRNFTGKTFVFD--GSQTGA--AKMQKSRSSNFIVKSPRPV--QGSYPI 213
Db 219 -----SYVGSMTANLLGSKYIWDKGVGVSGVKWVPLLSWVI--FTPTITTTGTSY-- 269
Qy 214 AHISYELNVLGSRGPRMRCITMDIPMSIVSRGVVASTSISSFSSRSRSPVFSHKLPR 273
Db 270 -----RRMTTL-----PKQOPMQKNKKVQ 291
Qy 274 SNSASCDSGNNLGDPPPLVLS-----NKAPRWHEQLRCWCLNF--HGRVTV---ASVKNF 323
Db 292 QASKLPDLWLENKEKIKQLCSRIHYNKISKQHE-----LDFDRGRGTGLRIQSSVKNF 345
Qy 324 QLVAVSDCEAGQTSERIILFGKVGKDMFTDYGYPISAFQAFICLSFETRIAC 379
Db 346 QLTLT-----ETPQTILQGRVDKARYVIDFRYPFGSYQAFCICLASIDSKLCC 395

RESULT 12
B86347
hypothetical protein F24J8.5 [imported] - *Arabidopsis thaliana*
C:Species: *Arabidopsis thaliana* (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: B86347
R:Theologias, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.; Li, J.H.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, R.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: B86347
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-360 <STO>
A:Cross-references: UNIPROT:Q9JPL4; UNIPARC:UPI00000A85E3; GB:AE005172; NID:g9454572; PII
C:Genetics:
A:Map position: 1

Query Match

Best Local Similarity 26.7%; Pred. No. 0.11; Mismatches 43; Conservative 21; Mismatches 50; Indels 47; Gaps 8;

QY 33 WSELPEELLREILIRVETVDGWPSSRRNVVACAGVCRSWRI-----LTK----- 77
Db 28 WKDIPVELLMRLSLVDD-----RNVIVASGCTGCRDAISFGLTRLSLSCNNNM 78
QY 78 ---EIVAVPEFSSKLTFFPISLKQSGPR--DSIVOCFIKRNRTQSVHLYGLTTSITDNG 132
Db 79 NSLVSLVPKFKVLOT--LNLKQDKPQLEDNAVEAIANHCHELQ--ELDLKSLKITDRS 134

QY 133 KFLLA-----ASKLRATCTDYIIISLRSDDISKRNNAYLGR 168
Db 135 LVLAHGCDFLTKNLSGCTSP-----SDTAIAYLTR 166

RESULT 13
D84634
hypothetical protein At2g24250 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: D84634
R:Lin, X.; Kaul, S.; Roundley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Mofatt, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: D84634
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-695 <STO>
A:Cross-references: UNIPROT:Q9ZUHO; UNIPARC:UPI000017A78D; GB:AE002093; NID:g4115385; P1
C:Genetics:
A:Gene: At2g24250
A:Map position: 2

Query Match 5.4%; Score 106.5; DB 2; Length 695;
Best Local Similarity 20.4%; Pred. No. 0.5; Mismatches 81; Conservative 52; Mismatches 153; Indels 111; Gaps 17;

QY 4 RSLQEMRSRPHRVHAAASTANSSDPFSWSSELPEELLREILIRVETVDGWPSSRRNVV 63
Db 309 QSAVKDLGESSMAMKKISTSSIMP-DWSQLPEELLHIISTHLE-----DHYFDVAV 360
QY 64 ACAGVCRSWR-----ILTKEIVAVPEP-----SS 87
Db 361 HARSVCRSWRSTPPPSLLRQSYSLPAPPLESKDLCTLEKVPFLFRVLTPPOAADAS 420
QY 88 KLTFFPISLKQSGPRD-----SLVOCFIKRN-RNTQ-----SVHLYGLT 125
Db 421 SEYFLGGLGQDKSNHVELPSPQLQSVKVNVPGETEPIILMMLDCQIIPIGHKYRLMIGN 480

QY 126 -----TSLTDNGKFLAASKLKRATCTDYIIISLRSD-----ISKRSNA----- 164
Db 481 PEEYSAAFLPLNEQGGGFVALLD-----CTDLFLVLRSTEMRWIRLEKSTASCKEL 534
QY 165 --YLGMRNRLGTFKFTVDGSGTGAARKMQKRSNFIKVSPPVQGSYPIAHISVELNV 222
Db 535 FTFRGRFYATFFNGDTFVIDPSSLEATPLTPHIDSNFL-----VPSGNEELFLV--KTDF 587

QY 223 LGSRGPRMRCIMDTIPMSIVSRGVVASTSISSF--SSRSSVPFRSHKPLRSNASC 280
Db 588 LRCKVRLDBEAAEWVSDGLGRVFLFGHLGNFYCSAKELP----HCGGLTGDSILPT 643

QY 281 DSGNNLGDPL--VLNKAQPRWHEOLRCWCLNFHGRV 315
Db 644 VGRNVTYPKYGVHTNKR-KAEDNINCWSSRENRV 679

RESULT 14
H85068
N7-like protein [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 31-Dec-2004
C:Accession: H85068
R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring Nature 402, 769-777, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:20083488; PMID:10617198
A:Accession: H85068
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-322 <STO>
A:Cross-references: UNIPROT:Q9MOU7; UNIPARC:UPI00000A04CC; GB:NC_001268; NID:g7267308; P1
C:Genetics:
A:Gene: AT4G05480
A:Map position: 4
C:Superfamily: F-box containing protein

Query Match 5.2%; Score 103; DB 2; Length 322;
Best Local Similarity 20.9%; Pred. No. 0.34; Mismatches 56; Conservative 54; Mismatches 104; Indels 54; Gaps 14;

QY 2 TFRSL-----QEMRSRPHRVHAAASTANSSDPFSWSSELPEELLREILIRVETVDGWP 57
Db 6 TLQSLLMKEDBEQRNK-RRTTSTMFLPKKDEERINWVDLPPELTSTILLSRLSVTDILD-N 63
QY 58 SRNVVACAGVCRSWRIITKE-----IVAVPEFSSKLTFFPISLKQSG----- 99
Db 64 ARK-----LCRRWRICKDPFPMWRKINLRDCLMYEFDPSMCHRIVDLSQGLLEINI 116
QY 100 ---PRDSLVOCFIKRNRTQSVHLYGLTSLTDNGKFLAASKLKRATCTDYIIISLRSD 156
Db 117 EHFVSDLSLIVDRSCNLKS--LGISIVEPMTNKG-----VMNGIEKLPLETLVIFHSS 170

QY 157 ---DISKRSNA--YLGMRNRLGTFKFTVDGSGTGAARKMQKRSNFIKVSPPVQGSY 211
Db 171 IKLDLKAIGHACPOLKTLKANSLSSEL-AHDISQVGYPTLLEC-DDDALAIASMPK--- 225

QY 212 PIAHISYELNVLGSRGPRRMRCIMDTIP 239
Db 226 -LRHLQLMGNGLTWTG---LNAILDGCP 249

RESULT 15
H85063
hypothetical protein AT4G05080 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 31-Dec-2004
C:Accession: H85063
R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring Nature 402, 769-777, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:20083488; PMID:10617198
A:Accession: H85063
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-391 <STO>
A:Cross-references: UNIPROT:Q9S9T0; UNIPARC:UPI00000A6B39; GB:NC_001268; NID:g7267367; P1
C:Genetics:
A:Gene: AT4G05080
A:Map position: 4
C:Superfamily: hypothetical protein containing F-box domain

Query Match 5.0%; Score 98.5; DB 2; Length 391;
Best Local Similarity 20.1%; Pred. No. 1.1; Mismatches 62; Conservative 49; Mismatches 116; Indels 81; Gaps 13;

QY 35 ELPEELLREILIRVETVDGWPSSRRNVVACAGVCRSWRIITKEIVAVPEFSSKLTFFIS 94
Db 6 DLTQDLVKEILSRVPTISLG-----AVRSTCKGNWALSKDRI----- 42

QY 95 LKQSGPRDSLVOCFIKRNRTQSVHLYLG-----LTTSLTDNGKFLAASKLKRA 144
Db 43 LCKAKPKQOFHQGFMLSDYRLSRMFRNISTGTFKENGEEFVNLSVKRIGNLANKVEISHMY 102

```
QY 145 TCTDYIISLRSDDISKRSNAYLGRMRSNPLGTYFTVFDGSGTCAAKQKRSRNSNFIKVSP 204
Db |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
103 YCGILLCVTTTRLVIMNPYLGIKWIQLKTE-TWY---STFCLRYDNNKNHKILRFLD 158
QY 205 RVPQGSYPPIAHISYELNVLGSRGPRMRMCIMDTIPMSIVESRGVAVASTSISFSSRSSPV 264
Db ||||| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
159 N-KQGSYEI----YDLKYSWRA-----FDVIPKWDIDDDG-----QSASVKGNTY 199
QY 265 FRSHSKPLRSNSASCSD-SGNNLG---DPPLVLSNKAPRWHEQLRCWCNLFHGRVTVASV 320
Db ||: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
200 FRTIDE--TPNLLICPFTAEFRFGKLLDPPFQ-----HGMMSLSWV 238
QY 321 KNFQLVAV 328
Db |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
239 REEKLVAL 246
```

Search completed: December 23, 2005, 23:28:33
Job time : 43 secs

THIS PAGE BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 26, 2005, 03:28:09 ; Search time 1104 Seconds

(without alignments)
8539.032 Million cell updates/sec

Title: US-10-763-042-20

Perfect score: 1140

Sequence: 1 agaggttcggaagttact.....aaaccagaattgcctgtgaa 1140

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.Main.*

- 1: /cgn2_5/prodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_5/prodata/1/pubpna/US08_PUBCOMB.seq.*
- 3: /cgn2_6/prodata/1/pubpna/US09A_PUBCOMB.seq.*
- 4: /cgn2_6/prodata/1/pubpna/US09B_PUBCOMB.seq.*
- 5: /cgn2_6/prodata/1/pubpna/US10A_PUBCOMB.seq.*
- 6: /cgn2_6/prodata/1/pubpna/US10B_PUBCOMB.seq.*
- 7: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq.*
- 8: /cgn2_6/prodata/1/pubpna/US10D_PUBCOMB.seq.*
- 9: /cgn2_6/prodata/1/pubpna/US10E_PUBCOMB.seq.*
- 10: /cgn2_6/prodata/1/pubpna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1140	100.0	1140	8	US-10-763-042-20
2	567.8	49.8	1140	8	US-10-763-042-22
3	396	34.7	396	3	US-09-770-423-879
4	367.8	32.3	1629	7	US-10-425-114-8367
5	347.6	30.5	2161	7	US-10-424-599-6651
6	342.6	30.1	1218	8	US-10-763-042-14
7	340.6	29.9	1908	7	US-10-424-599-141814
8	326	28.6	469	3	US-09-770-444-216
9	290.4	25.5	1910	7	US-10-437-963-90528
10	250.4	22.0	1182	8	US-10-763-042-13
11	229.2	20.1	1002	7	US-10-425-114-19311
12	228.6	20.1	1939	7	US-10-424-599-112097
13	220.4	19.3	2358	8	US-10-425-115-119315
14	220.2	19.3	1183	7	US-10-767-701-12426
15	219.8	19.3	2451	7	US-10-424-599-8380
16	210.2	18.4	1164	8	US-10-763-042-17
17	203.6	17.9	2181	7	US-10-424-599-105685
18	195.4	17.1	1335	8	US-10-763-042-21
19	195.4	17.1	1338	6	US-10-225-068-129
20	195.4	17.1	1338	7	US-10-374-780A-2895
21	195.4	17.1	1338	9	US-10-225-068-129
22	193	16.9	1137	8	US-10-763-042-18
23	192	16.8	1878	7	US-10-424-599-64212

24	184	16.1	1287	8	US-10-763-042-16	Sequence 16, Appl
25	174.4	15.3	1357	7	US-10-437-963-57823	Sequence 57823, A
26	169.8	14.9	1910	7	US-10-425-114-22666	Sequence 22666, A
27	169.8	14.9	1930	7	US-10-425-114-22784	Sequence 22784, A
28	169.6	14.9	2253	7	US-10-437-963-47524	Sequence 47524, A
29	165.2	14.5	520	7	US-10-424-599-9312	Sequence 9312, Ap
30	165	14.5	1740	7	US-10-437-963-19517	Sequence 19517, A
31	164.8	14.5	1414	7	US-10-425-114-29388	Sequence 29388, A
32	164.8	14.5	1578	7	US-10-424-599-97364	Sequence 97364, A
33	162	14.2	1926	7	US-10-425-114-26712	Sequence 26712, A
34	162	14.2	2154	8	US-10-425-115-17768	Sequence 17768, A
35	160.8	14.1	1750	7	US-10-425-114-30141	Sequence 30141, A
36	160.8	14.1	1826	7	US-10-437-963-50806	Sequence 50806, A
37	160.4	14.1	571	7	US-10-021-323-4660	Sequence 4660, Ap
38	160.4	14.1	1030	8	US-10-767-795-1188	Sequence 1188, Ap
39	160.4	14.1	2238	7	US-10-437-963-97195	Sequence 97195, A
40	157.8	13.8	707	7	US-10-424-599-6654	Sequence 6654, Ap
41	154.6	13.6	1959	7	US-10-437-963-25577	Sequence 25577, A
42	153	13.4	1369	8	US-10-425-115-35180	Sequence 35180, A
43	148.4	13.0	1365	8	US-10-763-042-12	Sequence 12, Appl
44	147.4	12.9	971	7	US-10-424-599-69809	Sequence 69809, A
45	146.4	12.8	1599	7	US-10-425-114-22378	Sequence 22378, A

ALIGNMENTS

RESULT 1

US-10-763-042-20

; Sequence 20, Application US/10763042

; Publication No. US20050014266A1

; GENERAL INFORMATION:

; APPLICANT: Shaw, Jai-Fu

; TITLE OF INVENTION: PLANT TUBBY-LIKE PROTEINS

; FILE REFERENCE: 08919-099001

; CURRENT APPLICATION NUMBER: US/10/763,042

; CURRENT FILING DATE: 2004-01-21

; PRIOR APPLICATION NUMBER: US 60/441,380

; PRIOR FILING DATE: 2004-01-21

; NUMBER OF SEQ ID NOS: 48

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 20

; TYPE: DNA

; ORGANISM: Arabidopsis sp.

; US-10-763-042-20

Query Match	100.0%	Score 1140;	DB 8;	Length 1140;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1140;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGACGTTCCGAAAGTTTACTCCAGGAAATGCGGTCTAGGCCACACCGTGTAGTTTCAGCC	60	
DB	1	ATGACGTTCCGAAAGTTTACTCCAGGAAATGCGGTCTAGGCCACACCGTGTAGTTTCAGCC	60	
QY	61	GCGGCTCAACCGCTAATAGTTAGTTCAGACCCCTTTTCAGTGTGCGAGCTCCCGAGGAGCTG	120	
DB	61	GCGGCTCAACCGCTAATAGTTAGTTCAGACCCCTTTTCAGTGTGCGAGCTCCCGAGGAGCTG	120	
QY	121	CTTAGAGAAATCTGATTTAGGTTGAGACTGTTGACGGCGGCGGATTTGGCGTCCGCGCGA	180	
DB	121	CTTAGAGAAATCTGATTTAGGTTGAGACTGTTGACGGCGGCGGATTTGGCGTCCGCGCGA	180	
QY	181	AACGTGGTGGCTTGTCGCGCGGTTTCTCTAGCTGAGGATTTCTACCAAGGAGATTGTA	240	
DB	181	AACGTGGTGGCTTGTCGCGCGGTTTCTCTAGCTGAGGATTTCTACCAAGGAGATTGTA	240	
QY	241	GCTGTTTCTGAATTTCTCTTAATTTGACTTTTCCCTATCTCCCTCAAGCAAGTCTGTGCTCA	300	
DB	241	GCTGTTTCTGAATTTCTCTTAATTTGACTTTTCCCTATCTCCCTCAAGCAAGTCTGTGCTCA	300	
QY	301	AGAGATTCTTAGTTCAATGCTTTTAAACGTAATCGAAATCTCAATCGTATCATCTC	360	

Db 301 AGAGATTCTAGTTCATAGCTTTATAAAGCGTAATCGAAATACATCAATCGTATCATCTC 360
QY 361 TATCTCGGATTAACACTCTTTTGACGGATTAACGGGAAGTTTCTTCTGCTGCTTCTTAAG 420
Db 361 TATCTCGGATTAACACTCTTTTGACGGATTAACGGGAAGTTTCTTCTGCTGCTTCTTAAG 420
QY 421 CTGAGCGCGCACTTGGCACTGATTAACATCATCTCTTTGCGTTCAGACCATATCTCAAG 480
Db 421 CTGAGCGCGCACTTGGCACTGATTAACATCATCTCTTTGCGTTCAGACCATATCTCAAG 480
QY 481 AGAAGCAACCGGTATCTTGGGAGATGAGATCGAACTTCTCTTGGAAACAAAATTCACGGTC 540
Db 481 AGAAGCAACCGGTATCTTGGGAGATGAGATCGAACTTCTCTTGGAAACAAAATTCACGGTC 540
QY 541 TTTGATGTTAGTACAGACCGGAGCAGCGAAGATGACGAGAGCGGCTCTTCTTAATTTTCATC 600
Db 541 TTTGATGTTAGTACAGACCGGAGCAGCGAAGATGACGAGAGCGGCTCTTCTTAATTTTCATC 600
QY 601 AAAGTTTCACTAGATTCTCTGAGGAAGTTTACCCCATCGCTCACATTTTCATACGAGTTA 660
Db 601 AAAGTTTCACTAGATTCTCTGAGGAAGTTTACCCCATCGCTCACATTTTCATACGAGTTA 660
QY 661 AACGCTCTTAGGCTCTCGGGACCGAGAAATGCGTTGCAATCATGATACAAATACCTATG 720
Db 661 AACGCTCTTAGGCTCTCGGGACCGAGAGATGCGTTGCAATCATGATACAAATACCTATG 720
QY 721 AGCATCGTGGAGTCGCGAGGATGATGCTTCAACATCCATAGCTCTTTTTCCAGTCGG 780
Db 721 AGCATCGTGGAGTCGCGAGGATGATGCTTCAACATCCATAGCTCTTTTTCCAGTCGG 780
QY 781 TCATCACCACTTTTAGTCTCACTCAAAACCAATGCGGAGTAATAGTCATCATGTCAGC 840
Db 781 TCATCACCACTTTTAGTCTCACTCAAAACCAATGCGGAGTAATAGTCATCATGTCAGC 840
QY 841 GACTCAGGCAACCACTCGGAGATCCACCAATGCTGCTGAGCAACAAAGCTCCACGGTGG 900
Db 841 GACTCAGGCAACCACTCGGAGATCCACCAATGCTGCTGAGCAACAAAGCTCCACGGTGG 900
QY 901 CATGAGCAGTTAGCTTGTGCTGCTTAAATTTCCATGCTGAGTCAAGTGGCTTCGGTT 960
Db 901 CATGAGCAGTTAGCTTGTGCTGCTTAAATTTCCATGCTGAGTCAAGTGGCTTCGGTT 960
QY 961 AAGAACTTTTACGTTTGGCAGTTAGTCACTGTGAAAGCAGGCGAGACATCTGAGAGATC 1020
Db 961 AAGAACTTTTACGTTTGGCAGTTAGTCACTGTGAAAGCAGGCGAGACATCTGAGAGATC 1020
QY 1021 ATACTCCAGTTTGGAAAGTTGGGAAGGACATGTTTACCATGATTTATGGATATCCGATT 1080
Db 1021 ATACTCCAGTTTGGAAAGTTGGGAAGGACATGTTTACCATGATTTATGGATATCCGATT 1080
QY 1081 TCTGCGTTTCAAGCGTTTGTCTATCTGCTGAGCAGTTTGGAAACCGAATTTGCTCTGAA 1140
Db 1081 TCTGCGTTTCAAGCGTTTGTCTATCTGCTGAGCAGTTTGGAAACCGAATTTGCTCTGAA 1140

RESULT 2

US-10-763-042-22
; Sequence 22, Application US/10763042
; Publication No. US20050014266A1
; GENERAL INFORMATION:
; APPLICANT: Shaw, Jai-Fu
; APPLICANT: Lai, Chia-Ping
; TITLE OF INVENTION: PLANT TUBBY-LIKE PROTEINS
; FILE REFERENCE: 08919-099001
; CURRENT APPLICATION NUMBER: US/10763,042
; CURRENT FILING DATE: 2004-01-21
; PRIOR APPLICATION NUMBER: US 60/441,380
; PRIOR FILING DATE: 2004-01-21
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 1140

; TYPE: DNA
; ORGANISM: Arabidopsis sp.
US-10-763-042-22
Query Match 49.8%; Score 567.8; DB 8; Length 1140;
Best Local Similarity 72.8%; Pred. No. 6.9e-183;
Matches 810; Conservative 0; Mismatches 252; Indels 51; Gaps 4;
QY 58 GCCCGCGCTCAACCGCTAATAGTTTCAGACCCCTTTCAGTGTGTCGAGCTCCCGGAGGAG 117
Db 49 GCTGCCGATTCACCTTCCTGTGTTCATCGCAAGATTAATCGTGTGTCAGAGATTCCTGAAGAG 108
QY 118 CTGCTTTAGAGAAATCCTGATTTAGGCTTGAAGCTGTTGACGGCGGCGATTTGGCCGTCGCGG 177
Db 109 CTTCTTTAGGAGATTCCTGATTCGTGTTGAGCGGCGGACGGTGGCGGATGGCCGTCACGA 168
QY 178 CGAAACGTTGGTGGCTTGTGCGGCGCTTGTCTGCTAGCTGAGGATTCCTCAACGAGGATTT 237
Db 169 CGCAGCGTGGTGGCTTGTGCGGCGCTTGTCTGCTGCTGCGGCTACTTTATGACGAAACC 228
QY 238 GTAGCTGTTCTGATTTCTCTTAATTTGACTTTTCCCTATCTCTCCCTCAAGCAGTCTGCT 297
Db 229 GTCTGTTGCTCCCTGAGATCTCTTAACTTGAATTTTCCCATCTCTCTCAAGCAGCTGCT 288
QY 298 CCAAGAGATTTCTTAGTTCAATGCTTTTAAACCGTAATCGAAATACCTCAATCGTATCAT 357
Db 289 CCAAGGATTTCACTGCTTCAATGCTTTTAAACCGTAATCGAAATACCTCAATCATCATCAT 348
QY 358 CTCTATCTCGGATTAACCTCTTTTGAACGGAATTAACGGGAAAGTTTCTTCTGCTGCTTCT 417
Db 349 CTCTATCTCGGATTAACCAACTCTTTAAACGGAATGATGGAAGTTTCTTCTGCTGCTGCT 408
QY 418 AGCTGAGAGCGGCAACTTGTGACTGATTAATCATCATCTCTTTGCGTTTCAGACATATCTCA 477
Db 409 AAGTTTGAAGCACACAACTTGTACCGATTAATCATTTATCTCTTTACGTTTCTGATGATGTCG 468
QY 478 AAGAGAGCAACGCTATCTTGGGAGATGAGATCGAACTCTCTTGGAAACAAAATTCACG 537
Db 469 AAGAGAGCAACGCTTATGTTGGCAAGTGAATCGAACTTCTTAGGAACGAAATTCAT 528
QY 538 GTCTTTGATGTTAGTCAG-----ACCGGAGCAGCGAAGATGCGAAGAGCGGCTCT 588
Db 529 GTCTTTGATGGAATCTGCTGCTTTCAACGGGAGCGCAAAAGTTGAGAAAGAGCGCATCT 588
QY 589 TCTAATTTTCATCAAGTTTCACTAGAGTTCTCTCAGGGAAGTTACCCCATCGCTCACATT 648
Db 589 TATATTCGCGCAAAAGTTTTCAGCAAAAGTTCTCTTTGGAAGTTTCTCTGCTGCTCATATC 648
QY 649 TCATACGAGTTTAAACGCTTTAGGCTCTCGGGGACCGAGAAAGATGCGTTTGCATCATGGAT 708
Db 649 ACATATGAGCTGATGTTCTTAGGATCCCGGGGACCAAGAAAGATGCAATGTCTTATGGAC 708
QY 709 ACAATACCTTATGAGCATCGTGGAGTCGCGGAGGATGATGTTTCAACATTCATTAAGCTCT 768
Db 709 ACAATACCTTACAGCACAAATGGAGCTCAAGGAGTAGCTTCAGAACCATCATGAGTTTCCC 768
QY 769 TTTTCCAGTCGCTCATCACCAGTCTTTAGTCTCTCAATCAAAACCATTTGGTGTGAGCAACAA 828
Db 769 TTACTCGGTACTCGGTCACCTTATCCAGGTCTCTAGTCAAAACCATTAACGAGTA----- 823
QY 829 GCATCATGTAGCGACTCAGGCAACCACTGGGAGATCCACCATTTGGTGTGAGCAACAA 888
Db 824 -----GCTCAAGCCACCTGAAAGAAACACCATTAGTGTGAGCAACAG 867
QY 889 GCTCCAAGGTCGATGAGCAGTTTACGTTCTGCTGCTTAAATTTTCAATCGTTCGAGTCACA 948
Db 868 ACACCAAGGTCGACGAGCAGCTAGCTGCTGCTGCTTGAATTTTCAATGCTGCTGCTGCTGCT 927
QY 949 GTGGCTTCGTTTAAAGAACTTTTCAAGTTTGGGCG-----TTAGTCACTGTGAGCA 999
Db 928 GTAGCGTCAGTGAAGAACTTTTCAAGTTCGTCGTCGAGGAGTAGTGTGGCAGTCGACG 987
QY 1000 GGCGAGACATCTGAGA-----GGATCATACTCCAGTTTGGGAAAGTTGGGAAG 1047

Db 633 TTATGAAAGTGGCTCCCAAGCGCTGGTGAATCTAAAGCAAGTTTACCCAGGTTCCCTA 692
QY 623 AGGGAAGTTACCCCATCGCTCAATTTCAACAGATTAAACGCTTTAGGCTCTCGGGGAC 682
Db 693 CAGGCAACTATCCAGTGGCCCAATTTTCAATGAATGAATGCTAGGCTCAAGGGGTC 752
QY 683 CGAGAGAAATCGGTGATCATGATGATACATACCTATGAGCATCGTGGAGTCCGAGGAG 742
Db 753 CTAGGAGAAATGATGTTGTGATGATGATTCATCTCCGCCACTGCAATTTGAACCTGGAGGTG 812
QY 743 TAGTAGCTTCAACATCCATAAGCTCTTTTTCAGTGGCTCATCCAGGCTTTTAGTGTCTC 802
Db 813 TAGCACCTACACAGAC--TGAGTTTCTCTTAACAACATAGACATGTTTCTTTCATCC 870
QY 803 ACTCAAAACATTTGGCCAGTAATAGTGCATCATGTAGCACTCAGG-----CAACAACC 856
Db 871 CTTTTTTTCGATCAAAATCAATCGTGGGAAATTTCCATGCTGGACCAATTTGGTTGATC 930
QY 857 TGGAGATCCACATTTGGTGTGAGCAACAAAGCTCCAGGTGGCATGAGCAGTTACGTT 916
Db 931 AAAAGGATGGGATGCTAGTGTGTAAGAAACAAAGCCCTAGGTGGCATGAGCAGCTGCAAT 990
QY 917 GCTGGTGTCTTAATTTCCATGTTGCTGAGTCACTAGTGGCTTCGGTTAAGAACTTTTCAGCTTG 976
Db 991 GTTGGTGCCTGAACCTTTCAATGAGCGGTGACAAATGCTCAGTTAAGAACTTTTCAGCTGG 1050
QY 977 TGGCAGTTAGTG---ACTGTGAAGCAGGCGCAGACATCTGAGAGGATCATCTCCAGTTTG 1033
Db 1051 TGGCTTCTGCAGAAACGGACCTGCTGGACCAAGAACACGATAAGATCATCTCCAAATTTG 1110
QY 1034 GGAAAGTTGGGAAGACATGTTTACCATGATATGATGATATCGGATTTCTGGTTTCAAG 1093
Db 1111 GAAAGTTGGGAAGGATTTGTTTACAATGATATACCGGTACCCATCTCGGCATTTTCAG 1170
QY 1094 CGTTTCTATCTGCTGAGCAGTTTGAACCAAGAAATGCTGTA 1140
Db 1171 CATTTGCAATCTGCTCAGCAGTTTGTGATACCAAGATTTGCTGTGAA 1217

RESULT 5

US-10-424-599-6651
; Sequence 6651, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 6651
; LENGTH: 2161
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_106014C.1
US-10-424-599-6651

Query Match 30.5%; Score 347.6; DB 7; Length 2161;
Best Local Similarity 63.2%; Pred. No. 3.2e-107;
Matches 676; Conservative 0; Mismatches 364; Indels 30; Gaps 8;
QY 95 GCTGGTGGAGCTCCCGAGGAGCTGCTTAGAGAAATCTCTGATAGGTTGAGCTGTTG 154
Db 387 GTTGGGCAACATGCGCCCAAGAGCTTCTCCGAGAGGTCCTCTCCGAAATCGAG--GCCT 443
QY 155 ACGGCGCGAGTTGCGCGTCCGCGGCAACGTTGCTGGCTTGTGCGGCTTTGTGCTAGCT 214

RESULT 6

US-10-763-042-14
; Sequence 14, Application US/10763042
; Publication No. US20050014266A1
; GENERAL INFORMATION:
; APPLICANT: Shaw, Jei-Fu

Db 444 CCGAGGACACGTGGCGCCCGCGGAAGAGCGTCTCTCTCGCGCGCGCTCTCCGCGAGCT 503
QY 215 GGAGGATTTCTCACCAGGAGATTTAGCTGTTCTCTGAATTTCTCTCTCTCTCTCTCTCTCT 274
Db 504 GGAGACATATCACCAGGAGCATTTGTCACAAACGCGCGAACTCTCTCTCTCTCTCTCTCT 563
QY 275 CTATCTCTCTCTCAAGCAGTCTGCTCCAGAGATTTCTCTAGTTCAATGCTTTTATAAACGTA 334
Db 564 CCAATTTCTGTTAAACAGCTGGCCCAAGGGAATCTCTCTAGGGTCTTTTATAAGCGCA 623
QY 335 ATCGAAATFACTCAATCGTATCATCTCTATCTCGGAATTAACCTCTTTTGAACGATAACG 394
Db 624 ACGGTTCCACCAACATCTATTTGTTTCTCAGTTTAAACAGTACGCTAGCTGAAGATG 683
QY 395 GGAAGTTTCTTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 454
Db 684 GGAAGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 743
QY 455 CTTTCGCTTCAGACATATCTCAAGAGAGCAACGCGTATCTCTGGGAGAAATGAGATCCA 514
Db 744 CTCCTGACGCGAGATGATGTCAGAGGAGCAACTCTCTATGTTGGGAACTTAAGATCAA 803
QY 515 ACTTCTTTGGAAACAAATTCACGGTCTTTGATGGTGTAGTCAAGAC---CGAGGAGCGAAGA 571
Db 804 ACTTCTTTGGAAACAAATTTCAATCTATGATAGCCAGCTGCTCTCATACAGGAGCAAGA 863
QY 572 TCCAGAGAGCGGCTCTTC-----TAATTTCTCAAGTTTTCACCTAGAGTTCTCT- 621
Db 864 TTATGAAAGTGGCTCTCAAGGCTGGTGAATCTAAAGCAAGTTTTCACCAAGGTTCTCTA 923
QY 622 CAGGGAAGTTACCCCATCGCTCACAATTTTATACAGAGTT--AAACGCTCTTAGGCTCTCGGG 679
Db 924 CAGGCTAATCTCAGTGGCCCATTTTCAATGATGATGATGATGATGATGATGATGATGATG 983
QY 680 GACCGAGAGATGCGTGTGCTCATCATGATACAAATACCTATGAGCATGCTGGAGTCCGAG 739
Db 984 GTCTTAGGAGATGCAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1043
QY 740 GAGTAGTAGTTCAACATCCATTAAGCTCTTTTTCAGTGG-----GTTCATCACCAAGTCT 793
Db 1044 GTGTAGCTCTCACAGACTGAGTTTCTCTTAAACAACATAGACATGTTTCTCTTCAATCC 1103
QY 794 TTAGGTCTCACTCAAAACCATTTGCGCAGTAATAGTGCATCATGTAGCAGCTCAGGCAACA 853
Db 1104 CTTTTTTTTCATCAAAATCAAAATCTGCGGAAATTTCAATGCTGAGCACTGTTGATC 1163
QY 854 ACCTGGGAGATCCACATTTGCTGAGCAACAAAGCTCCACGGTGGCATGAGCAGTTAC 913
Db 1164 AAAAG---GATGGGATGCTAGTGTGTTGAAACCAAGGCCCTTAGTGGCATGAGCAGCTGC 1220
QY 914 GTTGTGCTGCTTAAATTTTCATGCTGAGTGCACAGTGGCTTCGGTTAAGAACTTTTCAGC 973
Db 1221 AATGTTGCTGCTGAACTTTTCAATGAGCGGTTGACAAATGCTCTCAGTTAAAACTTTTCAG 1280
QY 974 TTGTGCGAGTTAGTG---ACTGTGAAGCAGGCGCAGACATCTGAGAGGATCATCTCCAGT 1030
Db 1281 TGGTGGCTTCTGCAAGAAACCGACCTGCTGGACCAAGCAACGATAGATCATCTTCAAT 1340
QY 1031 TTGGGAAAGTTGGGAAGGACATGTTTACCATGATTAAGATATCCGATTTCTGCGTTTC 1090
Db 1341 TTGGGAAAGTTGGGAAGGATTTGTTTACAATGATTAACCGGTACCTATCTCTCGCATTT 1400
QY 1091 AAGCGTTTCTATCTGCTGAGCAGTTTGAACCAAGAAATTTGCTGCTGTA 1140
Db 1401 AGGCATTTTGAATCTGCTCAGCAGTTTGTGATACCAAGATTTGCTGTGTA 1450

APPLICANT: Lai, Chia-Ping
; TITLE OF INVENTION: PLANT TUBBY-LIKE PROTEINS
; FILE REFERENCE: 08919-099001
; CURRENT APPLICATION NUMBER: US/10/763,042
; PRIOR FILING DATE: 2004-01-21
; PRIOR APPLICATION NUMBER: US 60/441,380
; PRIOR FILING DATE: 2004-01-21
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 1218
; TYPE: DNA
; ORGANISM: Arabidopsis sp.
US-10-763-042-14

Query Match 30.1%; Score 342.6; DB 8; Length 1218;

Best Local Similarity 59.8%; Pred. No. 1.2e-105;

Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;

QY 32 GGTCTAGGCGCACCCGCTGTAGTTTCAGCGCGCGCTCAACCGCTAATAGTTTCAGACCCCTT 91
DB 89 GATCCAGGCTCAACGCTGTGTTCAGGATCTTCTGCTGTGTGATCTTTCACAGCA 148
QY 92 TCAGCTGGTCGGAGCTCCCGGAGAGCTGCTTAGAGAAATCCCTGATTAGGGTTGAGACTG 151
DB 149 GCTGCTGGCTAGTATGCTCCCGAGCTCTGAGAGATGTTCTTAGAGATTTGAGCAAT 208
QY 152 TTGACGGCGGCGAATGGCGCTGCGCGGAAACGCTGTGGCTGTGCGCGGTTTTCGTA 211
DB 209 CCGAAGACAC---TTGCGCGCTAGGAAAAATGTTTCTTTCGCTGTGTGTCGAGCA 265
QY 212 GCTGGAGGATCTCACCAAGGAGATGTAGCTGCTCCTGAATCTCTCTAATTCATTT 271
DB 266 ACTGGCGGAAATCGTCAAGAGATGCTCAGAGTTCCTGAGCTTTCTAGCAACTCAGTT 325
QY 272 TCCTATCTCCCTCAAGAGTCTGCTGCTCAAGAGATCTCTAGTTCAATGCTTTATAAAAC 331
DB 326 TTCCTATCTCCCTCAAGAGCGGGTCTTAGAGATCACTGTTCAATGCTATATTATGA 385
QY 332 GTAATCGAATACTCAATCGTATCATCTCTATCTCGGATTA---CTACTCTTTGACGG 388
DB 386 GAAACCGCAGCAATCAACTACTATCTATACCTCGGGTTAAACCAAGCAGCTTCAATG 445
QY 389 ATAAAGGAGATTTCTTCTGCTGCTTCTAGCTGAAGCGCGCACTTCGACTGATTA 448
DB 446 ATGATGGAAGTTCTTCTTCTGCTCAAGAGGTTTCGAGGCGCACTTCGACTGACTA 505
QY 449 TCATCTCTTTGCTTCAGAGATATCTCAAGAGAGCAACGCTATCTTGGGAGATGA 508
DB 506 TCATCTCTTAACTCGGATGATGCTCTCGAGGAAGCAATACCTATATCGGAAAGCTTA 565
QY 509 GATCGAATCTTCTTGGAAACAAATTCACCGTCTTTGATGGTAGTCAGACCGGAGCAG--- 565
DB 566 GATCTAACTTTCTGGGACCAAGTTCTACTGTCTATGACGCTCAGCGCAGCAATCTCGAA 625
QY 566 -----CGAAGATGCAAGAGAGCGCTCTTCTAATTTTCATCAAGTTTCACTAGAG 616
DB 626 CTCAGGTTACCAAGCCGCTTCAAGCAGACTTCTCAGTTTGAACCAAGTGAGCGCCGAGAA 685
QY 617 TTCTCAGGAGATTACCCATCGCTCACAATTCATACGATTAACGTTTAAAGCTTTAGGCTCTC 676
DB 686 TTCCATCTGGCAATATCTCTGTAGCATACTCTCATATGAGCTTAAAGCTTTGGGTTCCA 745
QY 677 GGGGACCGGAGAAATGCGTTGATCATGATGATACATACTATGAGCAGCTGTGGAGTCGC 736
DB 746 GAGGACCGGAGGAGTGCAGTGTGTCATGATGCGCATCTCTGATCAGCTGTAGAACCTG 805
QY 737 GAGGAGTAGTGTTCACATCATCAATAGCTCTTTTTCAGTC-----GGTCAATCACAG 790
DB 806 GAGGAACAGCTCCAACTCAGACGGAATTTGTCTCATAGCAATCTTGATAGTTTCCCTCAT 865
QY 791 TCTTTAGTCTCACTCAAAACCAATTCGCGAGTAATAGTGCATCATGTAGGACTCAGGCA 850

DB 866 TCTCCTTTCTTCAAGTCGAAATCAATTCGTGTCAGAGAGTCTCCCTTCTGTCTCATCTG 925
QY 851 ACACCTGGAGATCCACCAATTTGGTCTGAGCAACAAAGCTCCACGGTGGCATGACAGT 910
DB 926 CTGCTCAGAAGGAGAGACTGCTTGTCTGAAACAAAGCGGCCAGATGGCAGCAACAGC 985
QY 911 TAGCTTGTGCTGTCTTAAATTTCCATGGTCGAGTCAAGTGGCTTCGGTTAAGAACTTTC 970
DB 986 TCCAGTGTGCTGTGCTCACTTCAATGGGAGAGTCAAGTGGCTTCGGTCAAAACTTTC 1045
QY 971 AGCTTGTGCGAGTTAGTGACTGTGAA---GCAAGGCGAGACATCTGAGAGATCATACTCC 1027
DB 1046 AGCTGTAGTCTGCTCTGAGAAATGGACCTGCGAGGACCTGAGCAGCAAAACGTTATCTCC 1105
QY 1028 AGTTTGGGAAAGTTGGGAGGACATGTTTACCATGATTTATGATATCCGATTTCTGCGT 1087
DB 1106 AGTTTGGAAAGTCGGAAGAGATGTTTCAATGGATTTATCAGTACCTTATCTCTGCT 1165
QY 1088 TTCAAGCGTTTGTCTATCTGCTGAGCAGTTTGTGAACACAGAAATTCCTGTGAA 1140
DB 1166 TCCAGGCTTCAACATTTGCTCAGCAGTTTTCGACACCAAGATAGCATGTGAA 1218

RESULT 7

US-10-424-599-141814
; Sequence 141814, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 141814
; LENGTH: 1908
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_9906C.1
US-10-424-599-141814

Query Match 29.9%; Score 340.6; DB 7; Length 1908;
Best Local Similarity 61.0%; Pred. No. 7.6e-105;
Matches 651; Conservative 0; Mismatches 389; Indels 27; Gaps 5;

QY 95 GCTGTGCGAGCTCCCGGAGAGCTGCTTAGAGAAATCTCGATAGGGTTGAGACTGTTG 154
DB 482 GTTGGCTTAACATGCCCCCTGAGCTTTTGAGGAGATGTCATGAGATTTGAGGCTCTG 541
QY 155 ACGGCGGCTTGGCGCTGCGCGCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 214
DB 542 AGATTTCC---TGGGCTGCCAGAAACATGTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 598
QY 215 CGAGGATTTCTCAAGAGGAGATTTGAGCTGTTCTTGAATTTCTTCTTAAATGACTTTCC 274
DB 599 GAGAGAAATCATGAGGAAATTTGCAAGTCCCTCACTATCCGCAAGCTGACATTC 658
QY 275 CTATCTCCCTCAAGCAGTCTGTCAGAGATTTCTTAGTTCAATGCTTTTATAAAGCTA 334
DB 659 CAATTTCTTGAAGCAGCTGTCAGAGGAGCTCTCTTCTTCAATGCTATATTAACGAA 718
QY 335 ATCGAATACTCAATGATATCTCTATCTCGGATTAACCTCTTTGACGAGTAACG 394
DB 719 ATCGCAGTAATCAACATATTTATCTGTTTCTTGGTTTAAACCAAGCTCAACTGATGAAG 778
QY 395 GGAAGTTTCTTCTGCTGCTTCTTAAGCTGAAGCGCGCAACTTGCACCTGATCATCTCT 454
DB 779 GCAAGTTCTTCTTCTGCAAGGAGGCTGCACTCAGCTGACTATATTTATCT 838

QY 455 CTTTGGTTTCAGACGATATCTCAAGAGAGCAACGCGTATCTTGGGAGATGAGATCGA 514
Db 839 CTCCTAACTGTGATGATGATCAAGAGGAGTAGTACCTATATGTAAGAGTTGAGATCAA 898
QY 515 ACTTCTCTGGAAACAAAATTCACGGTCTTTTGATG-----GTAGTCAGACCGGAG 562
Db 899 ACTTCTTGGCACCAAAATTCACAGTGTATGATGCACACCCCTCAAATTTATGAGGCCAAG 958
QY 563 CAGCGAAGATGCAAGAGACCGCTCTTCTAAATTTATCAAGTTTCACTAGAGTTCTCTC 622
Db 959 TTACAAAGTCTCGTTCCACAGGCTAGTTAGTCTCAAGCAAGTTTCTCCAAGAGTTCTCG 1018
QY 623 AGGAGAGTTACCCCATCGCTCAATTTTCATACGAGTTAAACGTTTAGGCTCTCGGGGAC 682
Db 1019 CTGGCAACTATCCCATTTGCTCATGTGTCTATGATCTGAATGTTTGGGCTCTAGGGGCC 1078
QY 683 CGAAGAAGATGGGTTGCATCATGATGATACATACCTATGAGCATCGTGGAGTCGCGAGGAG 742
Db 1079 CTAGATAATGCATGTTGATGGATGCCATCCCTGCCCTCAGCTGTTGAACCTGGAGGTG 1138
QY 743 TAGTAGCTTTCAACATCCATAAGCTCTTTTTCAG-----TCGTCATCACACAGTCTTTA 796
Db 1139 TGGCCCAACACAGACTCAATTTCTTCATAGCAGAAATGATATCTTCTCCATCCATCCCTT 1198
QY 797 GGTCTCACTCAAAACCATTTGGCAGTAATAGTGCATCATGTAGCGACTCAGGCAACAACC 856
Db 1199 TCTTTAGATCAAAATCAACCCGGATGGCAATCTCCCA---ACAGTACCTTTGACTTTGTC 1255
QY 857 TGGGAGATCCACCATTTGGTGTGAGCAACAAAGCTCCAGGTGGCATGAGCAGTTACGTT 916
Db 1256 AATATGAGGGGACACTGGTATTAGAAACAAGTCCCAAGTGGCATGAGCAACCTTCAT 1315
QY 917 GCTGTGCTTAAATTTCCATGTCGAGTCAAGTGGCTTCGGTTAAAGAACTTTACGCTTG 976
Db 1316 GCTGTGCTGTAACCTTCAATGGCGAGTGACAGTTGCTTCAGTTAAAAATTTCCAGCTGG 1375
QY 977 TGGCAGTTAGTACTGTGAAG---CAGGGCAGACATCTGAGAGGATCATACTCCAGTTTG 1033
Db 1376 TTGCTTCTCCCAAAAATGGAGTTTCTTGAGCAGGCTCAGGAAAATGTAATCTACAGTTTG 1435
QY 1034 GGAAGTTGGGAGGACATGTTTACCATGGATTTATGGATATCCGATTTCTGGGTTTCAAG 1093
Db 1436 GAAAAGTTGGAAGGATGATTTCACCATGGATTCAGTATCCAACTCTGCTTTCAG 1495
QY 1094 CGTTTGTCTATCTCCCTGAGCAGTTTTCAAAACCAAGATTCGCTGTGAA 1140
Db 1496 CATTTGCAATATGCTTAGCAGCTTTTGACACCAAGATTGCTTGTGAA 1542

RESULT 8
US-09-770-444-216/c
; Sequence 216, Application US/09770444
; Patent No. US2002023280A1
; GENERAL INFORMATION:
; APPLICANT: Gorlach, Jörn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Matthew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Garcia, Carlos A.
; APPLICANT: Krickler, Maja
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil

; APPLICANT: Hurban, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; FILE REFERENCE: 2027 (PARA-016PRV)
; CURRENT APPLICATION NUMBER: US/09/770,444
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,502
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 216
; LENGTH: 469
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-770-444-216

Query Match 28.6%; Score 326; DB 3; Length 469;
Best Local Similarity 100.0%; Pred. No. 3.3e-100;
Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 815 TCGCAGTAATAGTCATCATGTAGCGACTCAGGCAACAACCTGGGAGATCCACCATGG 874
Db 469 TCGCAGTAATAGTCATCATGTAGCGACTCAGGCAACAACCTGGGAGATCCACCATGG 410
QY 875 TCGTCAGCAACAAGCTCCACGGTGGCATGAGCAGTTACGTTGCTGCTTAAATTTCC 934
Db 409 TCGTCAGCAACAAGCTCCACGGTGGCATGAGCAGTTACGTTGCTGCTTAAATTTCC 350
QY 935 ATGTCGAGTCACAGTGGCTTCGGTTAAGAACTTTACGTTTGGCAGTTAGTACTGTG 994
Db 349 ATGTCGAGTCACAGTGGCTTCGGTTAAGAACTTTACGTTTGGCAGTTAGTACTGTG 290
QY 995 RAGCAGGGCAGACATCTCAGAGGATCATACTCCAGTTTGGGAAAGTTGGGAAGGACATG 1054
Db 289 RAGCAGGGCAGACATCTCAGAGGATCATACTCCAGTTTGGGAAAGTTGGGAAGGACATG 230
QY 1055 TTACCATGGATTTATGGATATCCGATTTCTGCGTTTCAAGCGTTTCTATCTCCCTGAGCA 1114
Db 229 TTACCATGGATTTATGGATATCCGATTTCTGCGTTTCAAGCGTTTCTATCTCCCTGAGCA 170
QY 1115 GTTTTGAACCAAGAAATTCGCTGTGAA 1140
Db 169 GTTTTGAACCAAGAAATTCGCTGTGAA 144

RESULT 9
US-10-437-963-90528
; Sequence 90528, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 90528
; LENGTH: 1910
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_8918C.1
US-10-437-963-90528

Query Match 25.5%; Score 290.4; DB 7; Length 1910;

	Best Local Similarity	58.5%; Pred. No. 1.2e-87;	Mismatches	611; Conservative	0; Mismatches	406; Indels	27; Gaps	5;
QY	93	CAGCTGTCGGAGCTCCCGGAGGAGTGTCTTAGAGAAATCCTGATTAGGCTTGAGACTGT	152					
Db	842	CTGCTGGCGCAGCTGCCGCCGAGCTGCTCCGGAGGTGCTGGTGAGGATCGAGGATC	901					
QY	153	TGACGGCGGCATTTGGCGTTCGGCGGAACAAGTGTGGTCTGTGTGCGCGCTTTGTGCTAG	212					
Db	902	GGAAGTGTG---GTGGCCGTTCGGCGGAGGACGTGTGGCGTGGCGCGGCTCTGCCGAG	958					
QY	213	CTGAGGATTTCTCACCAAGGAGATTGTAGCTGTTCCTGAATTTCTCCTCTAAATTGACTTT	272					
Db	959	CTGAGGGGATCACC AAGGAGATTCGTC CGCGTCCCGAGCGGTCCGCAAGCTCACCTT	1018					
QY	273	CCCTATCTCCCTCAAAGCAGTCTGGTCCAAGAGATTCTCTAGTTCAATGCTTTATAAAG	332					
Db	1019	CCCCATCTCGCTCAAGCAGCTGCCCAAGAGCGGCACCTCTTAATGTTTTCAAGAAG	1078					
QY	333	TAATCGAAATPACTCAATCGTATCATCTCTATCTCGGAATTAACCTCTTTGACGGATAA	392					
Db	1079	AAACCGGACTACTCAGACATATTATCTGTATATTGGGCTCACAGAAGCACTTGCCGATGA	1138					
QY	393	CGGGAGTTTTCTTCTGCTCTTCTAGCTGAAGCGGCAACTTGCACCTGATTACATCAT	452					
Db	1139	TGGGAGTTTCTTACTTGTCTGCA CGCAATCGAGGAAGCCCACATGCA CAGCTACCTGAT	1198					
QY	453	CTCTTTCGCTTCAGACGATATCTCAAGAGAAAGCAACGCGTATCTTGGGAGAAATGAGATC	512					
Db	1199	TTCCCTTGACATGAGTGATATGTCAAGGGAGGCAACACCTATATTGGCAAGCTTAGGTC	1258					
QY	513	GAACCTCTTTGGAAACAAAATTCACGGTCTTTGATG-----GTAGTCAGACCGG	560					
Db	1259	AAACTTTCTCGGAACAAAGTTTACCGTCTATGATGCCCATCCACCATATGATGGAGCTGT	1318					
QY	561	AGCAGCGAAGATGCAGAGAGCGCTCTTCTTAATTTTCATCAAGATTTTCACCTAGAGTTCC	620					
Db	1319	TGTTCTCAAAAGTCGGTCTGCA CGTGTGGTCTGAACCAAGTCTCCCGAGAGTTCC	1378					
QY	621	TCAGGGAAGTTACCCCATCGCTCACATTTTCATACGAGTAAACGCTCTTAGGCTCTCGGGG	680					
Db	1379	TGCCGGGAATTAACAGTTTTCACATATTTCTACGAATGAATGTTTTGGGTGCAAGAGG	1438					
QY	681	ACCGAGAAGATCGTTGCATCATGGATACAATA CCTATGAGCATCGTGAGTTCGCGAGG	740					
Db	1439	TCCAAGAAGGATGAAC TGCATATGGATTCCATCCCCACATCAGCTGTTCAAGAGGGTGG	1498					
QY	741	AGTAGTAGC--TTCAACATCCATAGCTCTTTTTCAGTCGGTCATCACAGCTCTTTAGG	798					
Db	1499	GAAGAGCTCCAACACAGACTGAATTTCTCTCAGTGGCTCGACTCTTTCCCATCGATTTTC	1558					
QY	799	TCTCTCACTAAAA CCATTGCCAGTAATAGTGCATCATGTAGCGACTCAGGCAACAACTTG	858					
Db	1559	ATTCTTCAGATCAAAATCACCTCGTATAGAGATAGTGCACATTTTACAAATTATCTACTCAG	1618					
QY	859	GGAGATCCACCATTTGGTGTCTGAGCAACAAAGCTCCAGGTGGCATGAGCAGTTACGTTGC	918					
Db	1619	AAGGAAGAAAGCTGGTTCCGAGAAATATCTCTTAGTGGCATGAGCAACTGCAATGA	1678					
QY	919	TGTTGCTTAATTTCCATGG-----TCGAGTCACTAGTGGCTTCGGTTTAAAGAACTTTTCAG	972					
Db	1679	TGTTGCTCTCAACTTCCTCGTGGACGCCACGGGCTACTCTTGCTCTGTGTGAAAAA ACTTTTCAG	1738					
QY	973	CTTTGTGCGAGTTAGTGACTGTGAGCAGGGCA----GACATCTGAGAGGATCATACTCCA	1028					
Db	1739	CTGTTGTCTGTATGAAAA TTGGACCAACTAAC CAGGAGCAGGCAAGGTGATTTCTCTCCA	1798					
QY	1029	GTTTGGGAAAGTTGGGAAGGACATGTTTTTCA CATCGAATTATGGATATCCGATTTCTCGGTT	1088					
Db	1799	GTTTGGAAAGATTGAGAAGGACTTGTTCACCATGSACTACCGTTATTCGATATCAGCAATT	1858					
QY	1089	TCAAGCGTTTGTCTATCTGCGCTGAG 1112						

Db 1859 TCAATCATTTGCCAATCTGTCTGAG 1882

RESULT 10
US-10-763-042-13
; Sequence 13, Application US/10763042
; Publication No. US20050014266A1
; GENERAL INFORMATION:
; APPLICANT: Shaw, Jai-Pu
; APPLICANT: Lai, Chia-Ping
; TITLE OF INVENTION: PLANT TUBBY-LIKE PROTEINS
; FILE REFERENCE: 08919-099001
; CURRENT APPLICATION NUMBER: US/10763,042
; CURRENT FILING DATE: 2004-01-21
; PRIOR APPLICATION NUMBER: US 60/441,380
; PRIOR FILING DATE: 2004-01-21
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 1182
; TYPE: DNA
; ORGANISM: Arabidopsis sp.
US-10-763-042-13

Query Match 22.0%; Score 250.4; DB 8; Length 1182;
Best Local Similarity 55.4%; Pred. No. 4.8e-74;
Matches 585; Conservative 0; Mismatches 441; Indels 30; Gaps 4;

Qy	97	TGCTCGGAGCTCCGGAGGAGCTGCTTAGAGAAATCCTGATTAGGGTTGAGACTGTTGCAC	156
Db	142	TGGGCTTCTTTGGCCCTGAGTTGCTTCATGACATATATCTGGAGGTTGAAGAGATGA-	200
Qy	157	GGCGGCGATTGGCGGTCGCGCGCAAAACGTGGTGGCTTTGTCGCGGCGTTTGTGCGTGG	216
Db	201	--GACAGCTTTGGCCCGCTCGAGCTGCCGTTGTCTCTTTGTGCTTCAGTATGTAATCATGG	258
Qy	217	AGGATTTCTACCAAGGAGATTGTAGCTGTTCCTGAAATCTCCCTCTAAATTCACATTTCCCT	276
Db	259	AGAGGAATCACTATGGAGATTGTGAGATCCCTGACGAGCTGTGGGAAGCTCACATTTTCCA	318
Qy	277	ATCTCCCTCAAGCAGTCTGCTCCAAGAGATTCTAGTTCAATGCTTTATAAAACGTAAT	336
Db	319	ATCTCATGGAACAGCGCGGCCCTCGAGACTTCCAATTCANIGTTTTATTAAGAGGAC	378
Qy	337	CGAATAACTCAATCGTATCATCTCTATCTCGGATTAACACTACCTTTTGAACGATAACGGG	396
Db	379	AGAGCAACAGCTACATACATCTCTATTATGTGTTTGATGCTTCGGAGACTGAGAACGAC	438
Qy	397	AAGTTTCTTGTGCTGCTTCTAAGCTGAAGCGGCGCACTTGCACTGATTACATCATCTCT	456
Db	439	AAACTGTTGTTAGCAGCAAGAAGATTAGAAAGCGCATGTCACAGACTTTTATAATCTCC	498
Qy	457	TTGCGTTTCAGACGATATCTCAAGAGAGCAACCGCTATCTTTGGGAGAAATGAGATCGAAC	516
Db	499	CTATCTGCAAGAACTTCTCAGGAGCAGCAGTACTTATGTTGGCAATTAAGGTTCTGGT	558
Qy	517	TTCTCTTGGAAACAAAATTCACGGTCTTTTG---ATGGTAGTCAGACCGGAGCAGCGAAGATG	573
Db	559	TTTTCTGGGAACCAAGTTTCAATATATGACAAACCAACAGCATCATCCACAGCACAGGCC	618
Qy	574	CAGAAGAGCGGCTCTTCTTAATTTTCATCAAAAGTTTTCACCTGAGTTCCTCAGGGAAGTTAC	633
Db	619	CAACCTTAACCGAAGACTCCACCCGAAACAAGCGGCTCTTAACTACCTACGAAATAGTCTCT	678
Qy	634	CCCATCGCTCACATTTTCATACGAGTTAAACGTCTTAGGCTCTCGGGGACCGAGAGAATG	693
Db	679	ACCGTAGGAACATTAACCTACGAGCTCAATGTTCTTTCGCAACAGGGGACCTTAGAAGATG	738
Qy	694	CGTTGTCATCTGGATACAAATACCTTATGAGCATCTGCGAGTCCGCGAGGAGTAGTAGTCTCA	753
Db	739	CAGTGGCGTATGGATTCTATACCCCTCTCTCTCTGTTATGTCGTAACCGTCAAGTATGTTCAA	798
Qy	754	ACATCCATAAGCTCTTTTTCAGTCGGGTCAATCAACAGTCTTTTATGGTCTCACTCAAAACCA	813

QY 912 ACCTTGGTGGTAAATTTTCAATGTCGAGTCACAGTGGCTTCGGTTAAGAACTTTCA 971
DB 1845 GCAGTGTGGTGGTCACTTCGGGGGGGTGTACCGTGGGTGAGTGAAGACTTTCA 1904
QY 972 GCTTGGGAGTGTAGTGAAGTGAAGCGGCGAGACATCT-----GAGAGGATCAT 1022
DB 1905 GCTGGTGGCTCTGTGGTATCTTCCCTAAACATCCCGCAGCGGAGGAGAGGTGAT 1964
QY 1023 ACTCCAGTTTGGGAAAGTTGGGAGACATGTTTACCATGATATATGGATATCCGATTTTC 1082
DB 1965 CTTCCAGTTTGGGAAAGTGAAGGATATATTTACCATGGACTACCGGTATCCCACTCTC 2024
QY 1083 TCGGTTTCAAGCGTTTGTCTATCTGCTGAGCAGTGTGAAACACAGAAATTCCTGTGAA 1140
DB 2025 GCGTTTCCAGGCTTTTGGCATCTGCTGACCGACTTCGACACCAACCGGCTTCGGA 2082

RESULT 14
US-10-767-701-12426
; Sequence 12426, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 12426
; LENGTH: 1183
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS9122_1
US-10-767-701-12426

Query Match 19.3%; Score 220.2; DB 7; Length 1183;
Best Local Similarity 62.3%; Pred. No. 1.1e-63;
Matches 454; Conservative 0; Mismatches 253; Indels 22; Gaps 6;

QY 428 GCGCAACTGCACTGATATACATCATCTCTTTGGGTTTCAGACGATATCTCAAGAGAGCA 487
DB 4 GCGCCACATGCAAGACTAGTAAATTTCTTTGATAAGGTTGATATGTCAAGGGGAAGCA 63
QY 488 ACGCGTATCTTGGGAGATGAGATCGAACTTCTTCGAAACAAATTCACGGTCTTTGATG 547
DB 64 GCACCTATATTTGGCAAGCTAAGATCAAACTTCTTCGAAACAAAGTTCACTGTCTATGATG 123
QY 548 GTAGTCAGAC- --CGGAGCAGCGAAGATGCAAGAGCGCGCTCTTC-----TAATTT 595
DB 124 CTCATCCACCATATGATGAGCTGTGTCTCTCAAGAGTGCCTGACGTTGGTTGGTT 183
QY 596 TCATCAAGTTTCACTAGAGTTCTTCAGGAGATTTACCCATCGCTGACATTTTATAG 555
DB 184 TGAACAGGTTCTCCCTAGAGTTCCGGCTGGGAATTTATCTGTTTTCACATATTTCTTACG 243
QY 656 AGTTAAACGTTTAGGCTCTCGGGGACCGAAGAAATGCGTTTCATCATGGATACAAATAC 715
DB 244 AGCTGAATGTTCTGGGCTCCAGAGGTCCAGAAAGGATGAATCTGTGTATGATTCATCC 303
QY 716 CTATGAGCATCGTGGAGTCGCGAGGAGTAGTAGCTTCAACATCCATTAAGCTCTTTTCCA 775
DB 304 CGGCATCAGCTGTTCGAGGAGGAGGAGAGCTCTTACACAGACTGAATTTCCATTTAGCA 363
QY 776 GTCCGTCATCACCAAGTCTTAGGTCTCATCTCAAAACCATTTGCGCATATATAGTGATCAT 835
DB 364 GTCTTGACTCTTTTCCCATCAATTTCCATTTCTTCAGATC-TAAATACATCGGATAGACGT 422
QY 836 GTAGCGACTCAGGCAACACCTCGGAGATCCACCATTTGGTGTGAGCAACAAAGCTCCAC 895

DB 423 TCAACATCGCAGTCATCGAGATCGGATA-----GGTTGGTGTGAAGCAACAAGTCTCTCTA 477
QY 896 GGTGCGATGAGCAGTTAC-GTTGCTGTGTCTTAAATTTCCATGTGCGAGTCACAGTGGCT 954
DB 478 GGTGCGATGAACAACCTGAGATGTTGGTGCCTGAAATTTCCGTGGACGGGTCACTGTGTCT 537
QY 955 TCGGTTAAGAACTTTTCAGCTTGTGGCAGTTAGTGACTGTGAAGCAGGGGAGACATCTGAG 1014
DB 538 TCTGTTAAAACTTTTCAGTTGGTGGCTTCTGATGACATGGACCGGAAACCAAGATAAT 597
QY 1015 AGGA---TCATACTCCAGTTTGGGAAAGTTGGGAAGGACATGTTTACCATGATTAAGA 1071
DB 598 AACAAAGTCAATCTTCCAAATTTGGAAGATTTGGAAGAACTTGTTCACCATGACTACCGT 657
QY 1072 TATCCGATTTTCGCGTTTCAAGCGTTTGTCTATCTGCTGAGCAGTGTGAAACCAAGAT 1131
DB 658 TATCCAATATCAGCATTTCAAGCTTTTCAATTTTGTCTGAGCAGTTTGTATACAAAAT 717
QY 1132 GCCTGTGAA 1140
DB 718 GCCTCGGA 726

RESULT 15
US-10-424-599-8380
; Sequence 8380, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 8380
; LENGTH: 2451
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_107577C.1
US-10-424-599-8380

Query Match 19.3%; Score 219.8; DB 7; Length 2451;
Best Local Similarity 56.0%; Pred. No. 2.3e-63;
Matches 593; Conservative 0; Mismatches 382; Indels 84; Gaps 6;

QY 166 TGGCGTTCGCGCGAACAAGTGGTGGCTTGTGCGGGGTTTGTGAGTGGAGGATCTC 225
DB 612 TGGCTGTCTCGAAAGCATGTTAGTCATGTGTGAGTGTGCGAGTCTTGGGGAATATG 671
QY 226 ACCAAGGAGTTGAGTGTCTCTGAATTTCTCTTAAATTTGACTTTCCCTATCTCCCTC 285
DB 672 TCAAGGACATCGTTAAGGCCAGAGTTTGTGGCAAACTTACATCTCCCTGTCTCTTG 731
QY 286 AAGCAGTCTGTGTCGAAGAGATTTCTCTAGTTCAATGCTTTTATAAAACGTAATCGAAATCT 345
DB 732 AAGCAGCTGGGCCACGGATGGAATCATTAATGTTTTATCAAAAGAGATAAATCTAAT 791
QY 346 CAATCGATCATCTCTATCTCGGATTA---CTACCTCTTTGACGGATTAACGGAGATT 402
DB 792 TTAACATACCACTATTTCTTTGTCTCAGCCCTGTCTTTGTTAGTTGAAAATGGAAAATTC 851
QY 403 CTTCTTGTGCTTCTTAAGCTGAAGCGCAACTTGCACGTATTACATCATCTCTTTGCGT 462
DB 852 CTCCTTTCTGCTAAGAGGACAAGGAGAACAACTTTACAGAGTACGTTATTTCCATGGAT 911
QY 463 TCAGACGATATCTCAAGAGAAAGCAACGCGTATCTTTGGAGAAATGAGATCGAACTTCCT 522

```
Db 912 GCTGACACATCTCTAGATCCAGTAACACTTACATTTGGAAGCTGAGATCAAAATTTTCTT 971
Qy 523 GGAACAAAATTCACGGT-----CTTTGATGGTAGTCAGACGGGA 561
Db 972 GGTACAAAGATTCAATTATATGCGATACACAGCCCTCCATACCTCTTCTGCCCATATATGCCCT 1031
Qy 562 GCAGCGAAGATGCAGAGAGCGCTCTTCTAAATTCATCAAAAGTTTCACCTAGAGTTCTT 621
Db 1032 CCTATGACTGGGAAGACGAGCCGTAGATTTATTCAAAAGGTCTCGCCTAAGGTCCCA 1091
Qy 622 CAGGGAAGTTACCCCATCGCTCACAATTTCATACAGTTAAACGTCTTAGGCTCTCGGGGA 681
Db 1092 TCTGGGAGTTACAACTAGCTCAGGTAAACATATGAATTAATGTGCTTGGAACTCGAGGC 1151
Qy 682 CCGAGAAGAAATGGTTGCTATCATGGATCAATACC-----TATGAGCA 724
Db 1152 CCAAGAAAGATGCATCGCTTATGTCATTAATACACAGCTTCAGCACTTGATGCAGCGGC 1211
Qy 725 TCGTGGAGTCGGGAGGAGTAGTAGTTCACATCCATAAG----CTCTTTTCCAGTCGG 780
Db 1212 ACTGCTCTGGCCAGCCAGAGCTTCTCTCGTCCCTTGGAGACTCGTTTCGGAGCATC 1271
Qy 781 TCATCACCAAGTCTTTAGGTCTCACTCAAAACCAATTCGCGAGTAATAGTCATCATGTAGC 840
Db 1272 TCGTTTTCAAAGTCTCTAGATCGTTCGATGAGTTCAGCAGTTCACGATTTTCTGAGATT 1331
Qy 841 GACTCAGGCAACAACTCGGAGATC-----CACCATTGGTGTGAGCAACAAA 888
Db 1332 GGGGAATCTATCATCGAGGATGATGATGGCAAGATGAGACCCCTTGGTCTGAAAAACAAG 1391
Qy 889 GCTCCACGGTGGCATGAGCAGTTACGTTGCTGGTGTAAATTTCCATGTCGAGTCACA 948
Db 1392 CCTCCAAGATGGCAAGCAGTTACAAATGTTGGTGCCTTAATTTCCGTGGACGAGTAACA 1451
Qy 949 GTGGCTTCGGTTAAGAACTTTCAAGCTTGTGGCAGTTAGTGACTGTGAAGCAGGGGCAGACA 1008
Db 1452 GTTGGCTCTGTTAAAAACTTTCAAGTTGATGTGTCGCCACCACCCCGCTGGTGCACT 1511
Qy 1009 TC-----TGAGAGGATCATCTCCAGTTTGGGAAAGTT 1041
Db 1512 ACACCATCTCAACCAAGCTCCACCGGAGCATGATAAAATCATCTTCAAATTTGGCAAAGTT 1571
Qy 1042 GGGAGGACATGTTTACCATGGATTATGGATATCCGATTTCTCGGTTTCAAGCGTTTGCT 1101
Db 1572 GGTAAAGACATGTTTCAACATGGATTATCGGTACCCCTTTATCTGCAATTCGAAGCTTTTGA 1631
Qy 1102 ATCTGCTCAGCAGTTTGTAAACCAAGAAATTCCTGTGAA 1140
Db 1632 ATATGCTTGAGCAGCTTTGACACCAAAATTTGGCTTTGAA 1670
```

Search completed: December 26, 2005, 07:12:36
Job time : 1110 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: December 26, 2005, 01:49:06 ; Search time 5217 Seconds
(without alignments)
10223.744 Million cell updates/sec

Title: US-10-763-042-20
Perfect score: 1140
Sequence: 1 agacgttccgaagttact.....aaaccagaattgcctgtgaa 1140

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_est3:*

4: gb_hic:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_est7:*

9: gb_gss1:*

10: gb_gss2:*

11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1140	100.0	1560	4	CNS0A4PF
2	1127.2	98.9	1261	4	US0A5VU
3	1090.8	95.7	1438	4	CNS0A5OU
4	1004.8	88.1	1535	4	CNS0A5IP
5	830.2	72.8	1036	5	US0A5VU
6	598.2	52.5	795	5	US0A5VU
7	592.8	52.0	597	2	US0A5VU
8	569.2	49.9	1281	4	CNS0A1DW
9	516	45.3	516	1	US0A5VU
10	466	40.9	637	5	US0A5VU
11	449	39.4	598	1	US0A5VU
12	422.6	37.1	704	5	US0A5VU
13	388	34.0	420	8	US0A5VU
14	333	29.2	728	5	US0A5VU
15	331.2	29.1	358	8	US0A5VU
16	319.6	28.0	982	7	US0A5VU
17	308	27.0	934	7	US0A5VU
18	300.8	26.4	650	6	US0A5VU
19	298.4	26.2	391	1	US0A5VU
20	292	25.6	509	1	US0A5VU
21	285.8	25.1	909	8	US0A5VU
22	282.4	24.8	904	8	US0A5VU

23	270.8	23.8	351	8	T44750
24	259.4	22.8	577	8	DN776642
25	259	22.7	694	8	CX297509
26	258	22.6	811	8	CX673843
27	254.2	22.3	667	5	US0A5VU
28	253.8	22.3	876	7	CX260993
29	253.6	22.2	837	7	US0A5VU
30	250.4	22.0	632	5	US0A5VU
31	248.8	21.8	614	2	US0A5VU
32	245.8	21.6	663	2	US0A5VU
33	245.8	21.6	666	6	US0A5VU
34	244.8	21.5	933	8	US0A5VU
35	240.8	21.1	779	7	US0A5VU
36	240.8	21.1	809	6	US0A5VU
37	240.2	21.1	578	8	US0A5VU
38	239	21.0	844	6	US0A5VU
39	235.6	20.7	739	5	US0A5VU
40	235	20.6	548	1	US0A5VU
41	232.8	20.4	726	7	US0A5VU
42	231	20.3	232	9	US0A5VU
43	228.8	20.1	713	8	US0A5VU
44	228.6	20.1	897	8	US0A5VU
45	227.4	19.9	1107	7	US0A5VU

ALIGNMENTS

RESULT 1
CNS0A4PF
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source

CNS0A4PF 1560 bp mRNA linear HTC 06-FEB-2004
Arabidopsis thaliana Full-length cDNA Complete sequence from clone
GSLT7B4ZA09 of Flowers and buds of strain col-0 of Arabidopsis
thaliana (thale cress).
BX822936
HTC; GSLT cDNA.
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
Unpublished
2 (bases 1 to 1560)
Genoscope.
Direct Submission
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen)
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URUG INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
length
http://www.genoscope.cns.fr/cgi-bin/ggb?source=Arabidopsis.
Location/Qualifiers
1. .1560
/organism="Arabidopsis thaliana"
/mol_type="mRNA"

		/db_xref="taxon:3702" /clone="GSLTFFB74ZA09" /tissue_type="flowers and buds" /ecotype="Col-0" /plasmid="pCMVSPORT_6" 1. .1560 /gene="At3g06380"	
gene			
ORIGIN			
Query Match		100.0%; Score 1140; DB 4; Length 1560;	
Best Local Similarity		100.0%; Pred. No. 0;	
Matches 1140; Conservative		0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	ATGACGTTCCGAAAGTTTACTCCAGGAAATGCGGTCTAGGCCACACCGTGTAGTTCACGCC	60
Db	261	ATGACGTTCCGAAAGTTTACTCCAGGAAATGCGGTCTAGGCCACACCGTGTAGTTCACGCC	320
QY	61	GCGCCTCAACCGCTAATAGTTTACAGACCCCTTTTCAGCTGGTCGGAGCTCCCGGAGAGCTG	120
Db	321	GCGCCTCAACCGCTAATAGTTTACAGACCCCTTTTCAGCTGGTCGGAGCTCCCGGAGAGCTG	380
QY	121	CTTAGAGAAATCTGATTAGGTTGAGACTTTGACGGCGGCGATTGGCCGTGCGGGCGA	180
Db	381	CTTAGAGAAATCTGATTAGGTTGAGACTTTGACGGCGGCGATTGGCCGTGCGGGCGA	440
QY	181	AACGTGGTGGCTTGTCGCGCGTTTCTGCTAGCTGGAGGATTTCTACCAAGGAGATTGTA	240
Db	441	AACGTGGTGGCTTGTCGCGCGTTTCTGCTAGCTGGAGGATTTCTACCAAGGAGATTGTA	500
QY	241	GCTGTTCTCGAATTCCTCTAAATGACTTTCCCTATCTCCCTCAAGCAGTCTGTGCTCA	300
Db	501	GCTGTTCTCGAATTCCTCTAAATGACTTTCCCTATCTCCCTCAAGCAGTCTGTGCTCA	560
QY	301	AGAGATTCTCTAGTTCAATGCTTTATAAAGTAATCGAAATCTCAATCGTATCATCTC	360
Db	561	AGAGATTCTCTAGTTCAATGCTTTATAAAGTAATCGAAATCTCAATCGTATCATCTC	620
QY	361	TATCTCGGATTAACCTCTTTGACGGATACGGGAAGTTTCTTCTGCTCTCTTAAG	420
Db	621	TATCTCGGATTAACCTCTTTGACGGATACGGGAAGTTTCTTCTGCTCTCTTAAG	680
QY	421	CTGAAGCGCGCACTTGCACTGATTACATCATCTCTTTGCGTTTCAGACGATATCTCAAG	480
Db	681	CTGAAGCGCGCACTTGCACTGATTACATCATCTCTTTGCGTTTCAGACGATATCTCAAG	740
QY	481	AGAAGCAACCGGTATCTTGGAGAAATGAGATCGAACTTCTCTTGGAAACAAAATTCAGCGTC	540
Db	741	AGAAGCAACCGGTATCTTGGAGAAATGAGATCGAACTTCTCTTGGAAACAAAATTCAGCGTC	800
QY	541	TTTGATGTAGTCAGACCGGACGAGATGCGAAGGCGCTCTTCTTAATTTTCATC	600
Db	801	TTTGATGTAGTCAGACCGGACGAGATGCGAAGGCGCTCTTCTTAATTTTCATC	860
QY	601	AAAGTTTCACTAGTGTCTCTCAGGAAAGTTTACCCCATCGCTCACATTTTCATACAGATTA	660
Db	861	AAAGTTTCACTAGTGTCTCTCAGGAAAGTTTACCCCATCGCTCACATTTTCATACAGATTA	920
QY	661	AACGTTTAGGCTCTCGGGACCGAGAAAGTTCGTTGCATCATGATGATACAAATACCTATG	720
Db	921	AACGTTTAGGCTCTCGGGACCGAGAAAGTTCGTTGCATCATGATGATACAAATACCTATG	980
QY	721	AGCATCGTGGAGTCGGAGAGTAGTGTCTCAACATCCATAGCTCTTTTCCAGTCGG	780
Db	981	AGCATCGTGGAGTCGGAGAGTAGTGTCTCAACATCCATAGCTCTTTTCCAGTCGG	1040
QY	781	TCATCACCACTCTTTAGGTCCTCACTCAAAACCATTCGGCAGTAATAGTGCATCATGTAGC	840
Db	1041	TCATCACCACTCTTTAGGTCCTCACTCAAAACCATTCGGCAGTAATAGTGCATCATGTAGC	1100
QY	841	GACTCAGGCAACAACTGGGAGATCCCAATGGTGTCTGAGCAACAAAGCTCCACGGTGG	900
Db	1101	GACTCAGGCAACAACTGGGAGATCCCAATGGTGTCTGAGCAACAAAGCTCCACGGTGG	1160
QY	901	CATGAGCAGTTACGTTGCTGGTCTTAATAATTTCCATGGTCGAGTCACAGTGCCTTCGGTT	960
Db	1161	CATGAGCAGTTACGTTGCTGGTCTTAATAATTTCCATGGTCGAGTCACAGTGCCTTCGGTT	1220
QY	961	AAGAACTTTACGTTGCTGGCAGTTAGTCACTGTGAAGCAGGGCAGACATCTCAGAGGATC	1020
Db	1221	AAGAACTTTACGTTGCTGGCAGTTAGTCACTGTGAAGCAGGGCAGACATCTCAGAGGATC	1280
QY	1021	ATACTCCAGTTTGGGAAAGTTGGGAAGGACATGTTTACCATGATTATCGATATCCGATT	1080
Db	1281	ATACTCCAGTTTGGGAAAGTTGGGAAGGACATGTTTACCATGATTATCGATATCCGATT	1340
QY	1081	TCGCGTTTCAAGCGTTTGCCTATCTGCTGAGCAGTTTGAACACAGAAATTCCTGTGAA	1140
Db	1341	TCGCGTTTCAAGCGTTTGCCTATCTGCTGAGCAGTTTGAACACAGAAATTCCTGTGAA	1400
RESULT 2			
CNS0A5VU		1261 bp mRNA linear HTC 06-FEB-2004	
LOCUS		Arabidopsis thaliana Full-length cDNA Complete sequence from clone	
DEFINITION		GSLTFFB1ZD02 of Flowers and buds of strain col-0 of Arabidopsis	
ACCESSION		thaliana (thale cress).	
VERSION		BX822204	
KEYWORDS		HTC; GSLT cDNA.	
SOURCE		Arabidopsis thaliana (thale cress)	
ORGANISM		Arabidopsis thaliana	
REFERENCE		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
AUTHORS		Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;	
TITLE		rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.	
JOURNAL		1 (bases 1 to 1261)	
REFERENCE		Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C.,	
AUTHORS		Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,	
TITLE		Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.	
JOURNAL		Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:	
COMMENT		A Combined Approach to Evaluate and Improve Arabidopsis Genome	
JOURNAL		Annotation	
REFERENCE		Unpublished	
AUTHORS		2 (bases 1 to 1261)	
JOURNAL		Genoscope.	
TITLE		Direct Submission	
JOURNAL		Genoscope.	
COMMENT		Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :	
JOURNAL		BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr	
COMMENT		- Web : www.genoscope.cns.fr)	
JOURNAL		The sequences are based on single pass reads.	
REFERENCE		Life Technologies (a division of Invitrogen)	
AUTHORS		full-length libraries construction : Temple G.	
TITLE		Genoscope members carried out sequencing and annotation : Castelli	
JOURNAL		V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,	
COMMENT		Schachter V., Weissenbach J., Salanoubat M.	
JOURNAL		URGV INRA : Clepet C., Caboche M.	
REFERENCE		Annotation is based on the June 2003 version of the Arabidopsis	
AUTHORS		genome released by MIPS (Munich Information center for Protein	
TITLE		Sequences). 5 prime and 3 prime are assembled with Phrap.	
JOURNAL		http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full	
COMMENT		length	
JOURNAL		http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.	
REFERENCE		Location/Qualifiers	
AUTHORS		1. .1261	
TITLE		/organism="Arabidopsis thaliana"	
JOURNAL		/mol_type="mRNA"	
COMMENT		/db_xref="taxon:3702"	
JOURNAL		/clone="GSLTFFB1ZD02"	
COMMENT		/tissue_type="Flowers and buds"	
JOURNAL		/ecotype="Col-0"	
COMMENT		/plasmid="pCMVSPORT_6"	
JOURNAL		1. .1261	
COMMENT		/gene="At3g06380"	
JOURNAL		gene	
REFERENCE		ORIGIN	
AUTHORS		Query Match	
TITLE		Best Local Similarity	
JOURNAL		98.9%; Score 1127.2; DB 4; Length 1261;	
COMMENT		99.3%; Pred. No. 0;	

Matches 1132; Conservative 0; Mismatches 8; Indels 0; Gaps 0;									
QY	1	ATGACGTTCCGAAATTTACTCCAGGAAATGCGGTCTAGGCCACACCGTGTAGTTTACAGCC	60						
DB	57	ATGACGTTCCGAAATTTACTCCAGGAAATGCGGTCTAGGCCACACCGTGTAGTTTACAGCC	116						
QY	61	GCGGCTCAACCGCTAATAGTTTACAGACCTTTTCAGCTGTGCGAGCTCCGCGAGAGCTG	120						
DB	117	GCGGCTCAACCGCTAATAGTTTACAGACCTTTTCAGCTGTGCGAGCTCCGCGAGAGCTG	176						
QY	121	CTTAGAGAAATCCTGATTAGGTTTGAGACTGTTGACGGCGGCGATTTGGCGTCCGCGCA	180						
DB	177	CTTAGAGAAATCCTGATTAGGTTTGAGACTGTTGACGGCGGCGATTTGGCGTCCGCGCA	236						
QY	181	AACGTGCTGCTGTGCGCGGTTTGTCTAGCTGTGAGGATTTCTACCAAGAGATTGTA	240						
DB	237	AACGTGCTGCTGTGCGCGGTTTGTCTAGCTGTGAGGATTTCTACCAAGAGATTGTA	296						
QY	241	GCTGTTCTCGAATCTCTCAATTAATGACTTTCCCTATCTCCCTCAAGCAGTCTGTCCA	300						
DB	297	GCTGTTCTCGAATCTCTCAATTAATGACTTTCCCTATCTCCCTCAAGCAGTCTGTCCA	356						
QY	301	AGAGATTCTCTAGTTCAATGCTTTATAAAGCTAATCGAAATACTCAATCGTATCATCTC	360						
DB	357	AGAGATTCTCTAGTTCAATGCTTTATAAAGCTAATCGAAATACTCAATCGTATCATCTC	416						
QY	361	TATCTCGAATTAACCTCTTTTGACGGATAACGGGAAGTTTCTTCTGCTGTTCTAAG	420						
DB	417	TATCTCGAATTAACCTCTTTTGACGGATAACGGGAAGTTTCTTCTGCTGTTCTAAG	476						
QY	421	CTGAGCGCGCAACTGCACTGATTACATCATCTCTTTCGGTTCAGCATATCTCAAG	480						
DB	477	CTGAGCGCGCAACTGCACTGATTACATCATCTCTTTCGGTTCAGCATATCTCAAG	536						
QY	481	AGAAGCAACGCGTATCTTGGGAGATGAGATCGAACTTCTCTTGGAAACAAATTCAGGTC	540						
DB	537	AGAAGCAACGCGTATCTTGGGAGATGAGATCGAACTTCTTGGAAACAAATTCAGGTC	596						
QY	541	TTTGATGTAGTCAGACCGGAGCGAAGATGCGAAGAGCGGCTCTTCTAATTTATC	600						
DB	597	TTTGATGTAGTCAGACCGGAGCGAAGATGCGAAGAGCGGCTCTTCTAATTTATC	656						
QY	601	AAAGTTTACCTAGATTCTCAGGAAAGTTACCCATCGCTCACATTTTACACAGTTA	660						
DB	657	AAAGTTTACCTAGATTCTCAGGAAAGTTACCCATCGCTCACATTTTACACAGTTA	716						
QY	661	AACGCTCTTAGGCTCTCGGGACCGAAGAAATGCGTTGTCATCATGATCAATACCTATG	720						
DB	717	AACGCTCTTAGGCTCTCGGGACCGAAGAAATGCGTTGTCATCATGATCAATACCTATG	776						
QY	721	AGCATCGTGGAGTCGCGAGGATGATGCTTCAACATCCATAGCTCTTTTTCCAGTCGG	780						
DB	777	AGCATCGTGGAGTCGCGAGGATGATGCTTCAACATCCATAGCTCTTTTTCCAGTCGG	836						
QY	781	TCATCAGCATCTTTAGTCTCACTCAAAACATTTGCGCAGTAATAGTGCATCATGTAGC	840						
DB	837	TCATCAGCATCTTTAGTCTCACTCAAAACATTTGCGCAGTAATAGTGCATCATGTAGC	896						
QY	841	GACTCAGGCAACAACTGGGAGATCCACCATTTGGTCTGAGCAACAAAGCTCCAGCGCG	900						
DB	897	GACTCAGGCAACAACTGGGAGATCCACCATTTGGTCTGAGCAACAAAGCTCCAGCGCG	956						
QY	901	CATGAGCAGTTACGTTGCTGCTTAAATTTCCATGTCGAGTCAACAGTGGCTTCGGTT	960						
DB	957	CATGAGCAGTTACGTTGCTGCTTAAATTTCCATGTCGAGTCAACAGTGGCTTCGGTT	1016						
QY	961	AAGAACTTTTACGTTGTGAGTTTGTGATCTGTGAAGAGCGGAGATCTGTGAGGATC	1020						
DB	1017	AAGAACTTTTACGTTGTGAGTTTGTGATCTGTGAAGAGCGGAGATCTGTGAGGATC	1076						
QY	1021	ATACTCCAGTTTGGGAAGTTGGGAGGACATGTTTACCATGGATTTATGGATATCCGATT	1080						
DB	1077	ATACTCCAGTTTGGGAGAGTTGGGAAGGACATGTTTACCATGGATTTATGGATATCCGATT	1136						

QY	1081	TCCTGCTTTCAAGCGTTTCTCTATCTGCTGAGCAGTTTGAACACGAAATTCCTGTGAA	1140						
DB	1137	TCCTGCTTTCAAGCGTTTCTCTATCTGCTGAGCAGCTTTGAACACCAATTCCTGTGAA	1196						
RESULT 3									
CNS0A5QU		1438 bp	mRNA	linear	HTC 06-FEB-2004				
LOCUS	Arabidopsis thaliana Full-length cDNA complete sequence from clone								
DEFINITION	GSLTPGH27D03 of Hormone Treated Callus of strain col-0 of								
	Arabidopsis thaliana (thale cress).								
ACCESSION	BX824160								
VERSION	BX824160.1 GI:42464062								
KEYWORDS	HTC; GSLT cDNA.								
SOURCE	Arabidopsis thaliana (thale cress)								
ORGANISM	Arabidopsis thaliana								
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.								
REFERENCE	1 (bases 1 to 1438)								
AUTHORS	Castelli, V., Aury, J.M., Jallou, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, P., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.								
TITLE	Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation								
JOURNAL	Unpublished								
REFERENCE	2 (bases 1 to 1438)								
AUTHORS	Genoscope.								
JOURNAL	Direct Submission								
TITLE	Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :								
	BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr								
	- Web : www.genoscope.cns.fr)								
COMMENT	The sequences are based on single pass reads. Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G. Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jallou O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M. URVU INRA : Clepet C., Caboche M. Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap. http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full length								
FEATURES	http://www.genoscope.cns.fr/cgi-bin/ggb?source=Arabidopsis.								
source	Location/Qualifiers								
	1..1438								
	/organism="Arabidopsis thaliana"								
	/mol_type="mRNA"								
	/db_xref="taxon:3702"								
	/clone="GSLTPGH27D03"								
	/tissue_type="Hormone Treated Callus"								
	/ecotype="Col-0"								
gene	/plasmid="pCMVSPORT_6"								
	1..1438								
	/gene="At3g06380"								
ORIGIN									
Query Match	95.7%; Score 1090.8; DB 4; Length 1438;								
Best Local Similarity	99.5%; Pred. No. 0;								
Matches 1136; Conservative	0; Mismatches 2; Indels 4; Gaps 4;								
QY	1	ATGACGTTCCGAAATTTACTCCAGGAAATGCGGTCTAGGCCACACCGTGTAGTTTACAGCC	60						
DB	198	ATGACGTTCCGAAATTTACTCCAGGAAATGCGGTCTAGGCCACACCGTGTAGTTTACAGCC	257						
QY	61	GCGGCTCAACCGCTAATAGTTTACAGACCTTTTCAGCTGTGCGAGCTCCGCGAGAGCTG	120						
DB	258	GCGGCTCAACCGCTAATAGTTTACAGACCTTTTCAGCTGTGCGAGCTCCGCGAGAGCTG	317						
QY	121	CTTAGAGAAATCCTGATTAGGTTTGAGACTGTTGACGGCGGCGATTTGGCGTCCGCGCA	180						

Db	318		CTTAGAGAAATCCTGATTAGGTTGAGACTGTTGACGGCGGCGATTGGCCGTCGGCGA	377
Qy	181		AACGTGGTGGCTTGTGCGCGGCTTGTGCTAGCTGAGGATTCACCAAGAGATTGTA	240
Db	378		AACGTGGTGGCTTGTGCGCGGCTTGTGCTAGCTGAGGATTCACCAAGAGATTGTA	437
Qy	241		GCTGTTCTCGAATCTCTCTAAATTTGACTTTCCCTATCTCCCTCAAGCAGTCTGGTCCA	300
Db	438		GCTGTTCTCGAATCTCTCTAAATTTGACTTTCCCTATCTCCCTCAAGCAGTCTGGTCCA	497
Qy	301		AGAGATTCTCTAGTTCAATGCTTTATATAAACGTAATCGAAATPACTCAATCGTATCATCTC	360
Db	498		AGAGATTCTCTAGTTCAATGCTTTATATAAACGTAATCGAAATPACTCAATCGTATCATCTC	557
Qy	361		TATCTCGGATTAATACCTCTTTGACGGATTAAGGGAAGTTCTTCTGCTGCTCTTAAG	420
Db	558		TATCTCGGATTAATACCTCTTTGACGGATTAAGGGAAGTTCTTCTGCTGCTCTTAAG	617
Qy	421		CTGAAGCGCGCAACTTGCACTGATTACATCATCTCTTTGCGTTTCAGACGATATCTCAAG	480
Db	618		CTGAAGCGCGCAACTTGCACTGATTACATCATCTCTTTGCGTTTCAGACGATATCTCAAG	677
Qy	481		AGAAGCAACGCGTATCTTGGGAGATGAGATCGAACTTCTTTGGAAACAAATTCACGGTC	540
Db	678		AGAAGCAACGCGTATCTTGGGAGATGAGATCGAACTTCTTTGGAAACAAATTCACGGTC	737
Qy	541		TTTGATGTAGTCAGACCGGAGGAGGATGCGAAGAGCGGCTCTCTTAATTCATC	600
Db	738		TTTGATGTAGTCAGACCGGAGGAGGATGCGAAGAGCGGCTCTCTTAATTCATC	797
Qy	601		AAAGTTTACCTAGAGTTCTCTCAGGGAAGTATCCCATCGCTCACATTTTCATACAGTTA	660
Db	798		AAAGTTTACCTAGAGTTCTCTCAGGGAAGTATCCCATCGCTCACATTTTCATACAGTTA	857
Qy	661		AACGCTTAGGCTCTCGGG-ACCAGAGAAGATGCGTTGCA-TCATGGATACAAATACCTTA	718
Db	858		AACGCTTAGGCTCTCGGG-ACCAGAGAAGATGCGTTGCA-TCATGGATACAAATACCTTA	917
Qy	719		TGAGCATCTGAGTCCGAGGAGTAGTAGTTCAACATCCATAGCTCTTTTTCAGTTC	778
Db	918		TGAGCATCTGAGTCCGAGGAGTAGTAGTTCAACATCCATAGCTC-TTTTTCAGTTC	976
Qy	779		GCTATCACCACTTTTAGTCTCACTCAAACCAATGCGCAGTAAATAGTCATCATGTA	838
Db	977		GCTATCACCACTTTTAGTCTCACTCAAACCAATGCGCAGTAAATAGTCATCATGTA	1036
Qy	839		GCGACTTCAGGCAACACCTGGGAGATCCACCATTTGCTGAGCAACAAAGCTCCACCGT	898
Db	1037		GCGACTTCAGGCAACACCTGGGAGATCCACCATTTGCTGAGCAACAAAGCTCCACCGT	1096
Qy	899		GGCATGAGCAGTTACGTTGCTGGTGTAAATTTTCATGCTGAGTCACAGTGGCTTCGG	958
Db	1097		GGCATGAGCAGTTACGTTGCTGGTGTAAATTTTCATGCTGAGTCACAGTGGCTTCGG	1156
Qy	959		TTAAGAACTTTCAGCTTGTGGCAGTTAGTAGCTGTGAAGCAGGCGAGACATCTGAGAGGA	1018
Db	1157		TTAAGAACTTTCAGCTTGTGGCAGTTAGTAGCTGTGAAGCAGGCGAGACATCTGAGAGGA	1216
Qy	1019		TCATATCTCCAGTTTGGGAAGTTGGGAAGACATGTTTACCATGGAATTTGATATCCGA	1078
Db	1217		TCATATCTCCAGTTTGGGAAGTTGGGAAGACATGTTTACCATGGAATTTGATATCCGA	1276
Qy	1079		TTTCTGCGTTTCAAGCGTTTGTCTGCTGAGCAGGATTTGAAACAGAAATTCGCTGTG	1138
Db	1277		TTTCTGCGTTTCAAGCGTTTGTCTGCTGAGCAGGATTTGAAACAGAAATTCGCTGTG	1335
Qy	1139	AA	1140	
Db	1336	AA	1337	

CNS0A51P	1535 bp	mRNA	linear	HTC 06-FEB-2004
Arabidopsis thaliana Full-length cDNA Complete sequence from clone				
GSITPGH44ZH11 of Hormone Treated Callus of strain col-0 of				
Arabidopsis thaliana (Chale cress).				
Accession	EX824384			
Version	1	GI:42463622		
Keywords	HTC; GSLT cDNA.			
Source	Arabidopsis thaliana (chale cress)			
Organism	Arabidopsis thaliana			
Reference	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.			
Authors	Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.			
Title	Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation			
Journal	Unpublished			
Reference	2 (bases 1 to 1535)			
Authors	Direct Submission			
Journal	Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr - Web : www.genoscope.cns.fr)			
Comment	The sequences are based on single pass reads. Life Technologies (a division of Invitrogen) full-length libraries construction : Temple G. Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M. URGV INRA : Clepet C., Caboche M. Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences) . 5 prime and 3 prime are assembled with Phrap. http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full length			
Features	Location/Qualifiers			
source	1..1535			
	/organism="Arabidopsis thaliana"			
	/mol_type="mRNA"			
	/db_xref="taxon:3702"			
	/clone="GSITPGH44ZH11"			
	/tissue_type="Hormone Treated Callus"			
	/ecotype="Col-0"			
gene	1..1535			
	/plasmid="pCMVSPORT_6"			
	/gene="At3g06380"			
Origin				
Query Match	88.1%;	Score 1004.8;	DB 4;	Length 1535;
Best Local Similarity	94.1%;	Pred. No. 1.2e-305;		
Matches 1078;	Conservative 0;	Mismatches 62;	Indels 6;	Gaps 3;
Qy	1	ATGACGTTCCGAGTTTACTCCAGGAATCGGTTAGCCACACCGTGTAGTTCACGCC	60	
Db	200	ATGACGTTCCGAGTTTACTCCAGGAGATGCGGTTAGGCCACACCGTGTAGTTCACGCC	259	
Qy	61	GCGGCTCAACCGCTAAATAGTTTCAGTGGTGGAGCTCCCGGAGGAGCTG	120	
Db	260	GCGGCTCAACCGCTAAATAGTTTCAGTGGTGGAGCTCCCGGAGGAGCTG	319	
Qy	121	CTTAGAGAAATCCTGATTAGGTTGAGACTGTTGACGCGCGAATTCG- CGTCCGCGG	179	
Db	320	CTTAGAGAGATGCTGATTAGGTTGAGACTGTTGACGCGCGGATTCGCGCGG	379	
Qy	180	AAACGTGGTGGTTGTGCGCGGTTTGTTCGTAGCTGGAGATTCTCACCAGGAGTTGT	239	
Db	380	AAACGTGGTGGTTGTGCGCGGTTTGTTCGTAGCTGGAGATTCTCACCAGGAGATTGT	439	

Db 593 CTGAAGCGGCAACTTGACTGATATACATCATCTCTTTGCGTTACAGCATATCTCAAG 652
Qy 481 AGAAGCAACGGCTATCTTGGGAGATGAGATCGAACTTCTTGGGAGCAAAAATTCACGGTC 540
Db 653 AGAAGCAACGGCTATCTTGGGAGATGAGATCGAACTTCTTGGGAGCAAAAATTCACGGTC 712
Qy 541 TTTGATGTTAGTACAGCCGAGCAGGAGATGCGAGAGCGGCTCTTCTAATTTTCATC 600
Db 713 TTTGATGTTAGTACAGCCGAGCAGGAGATGCGAGAGCGGCTCTTCTAATTTTCATC 772
Qy 601 AAAGTTTACCTAGAGTTCTCAGGGAAGTTACCCCATCGCTCACATTTTCATACGAGTTA 660
Db 773 AAAGTTTACCTAGAGTTCTCAGGGAAGTTACCCCATCGCTCACATTTTCATACGAGTTA 832
Qy 661 AACGCTTTAGGCTCTCGGGGACGAGAAATGCGTTGTCATCATGATACAAATCCTATG 720
Db 833 AACGCTTTAGGCTCTCGGGGA-CGAGAGAAATGCGTTGTCATCATGATACAAATCCTATG 891
Qy 721 AGCATCGTGGAGTCGGGAGGATGAGCTTCAACATCCATAGCTCTTTTCCAGTCCG 780
Db 892 AGCATCGTGGAGTCGGGAGGATGAGCTTCAACATCCATAGCTCTTTTCCAGTCCG 951
Qy 781 TCATCACCAAGCTTTAGGTCCTCACTCAAAACCATTCGCGAGTAATAGTGCATCATGTAGC 840
Db 952 TCATCACCAAGCTTTAGGTCCTCACTCAAAACCATTCGCG-GTAAATAGTCATCATGTAGC 1010
Qy 841 GACTCAGGCAACAACTGG 859
Db 1011 GACTCAGGCAAAACTGG 1029

RESULT 6
BX837992
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
BX837992 795 bp mRNA linear EST 11-FEB-2004
BX837992 Arabidopsis thaliana cdna clone GSLTPGH21ZG06 5PRIM, mRNA sequence.
Arabidopsis thaliana cdna clone GSLTPGH21ZG06 5PRIM, mRNA sequence.
BX837992
BX837992.1 GI:42532075
EST.
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 795)
Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C.,
Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,
Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.
Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
Unpublished (2004)
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen)
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Clepet C., Caboche M.
URUG INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences).
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/EST
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.
Location/Qualifiers
1. .795
/organism="Arabidopsis thaliana"
/mol_type="mRNA"

/ecotype="Col-0"
/db_xref="taxon:3702"
/clone="GSLTPGH21ZG06"
/cissue_type="Hormone Treated Callus"
/clone_lib="Arabidopsis thaliana Hormone Treated Callus
Col-0"

ORIGIN

Query Match 52.5%; Score 598.2; DB 5; Length 795;
Best Local Similarity 95.6%; Pred. No. 4e-177;
Matches 626; Conservative 0; Mismatches 28; Indels 1; Gaps 1;

Qy 1 ATGAGCTTCCGAAGTTTACTCCAGGAATGCGGTCTAGGCCACACCGTGTAGTTCACGCC 60
Db 141 ATGAGCTTCCGAAGTTTACTCCAGGAATGCGGTCTAGGCCACACCGTGTAGTTCACGCC 200
Qy 61 GCCGCTCAACCGCTTAATAGTTTCAGACCCCTTTACGCTGGTGGAGCTCCCGGAGAGCTG 120
Db 201 GCCGCTCAACCGCTTAATAGTTTCAGACCCCTTTACGCTGGTGGAGCTCCCGGAGAGCTG 260
Qy 121 CTTAGAGAAATCCTGATTTAGGTTGAGACTGTTGACGGGGCGGATTTGGCGTCGGCGGA 180
Db 261 CTTAGAGAAATCCTGATTTAGGTTGAGACTGTTGACGGGGCGGATTTGGCGTCGGCGGA 320
Qy 181 AACGTGGTGGCTTGTGCGCGGCTTTGTCTAGCTGAGGATTTCTCAACAGGAGATTGTA 240
Db 321 AACGTGGTGGCTTGTGCGCGGCTTTGTCTAGCTGAGGATTTCTCAACAGGAGATTGTA 380
Qy 241 GCTGTTCTCGAATTCCTCTTAAATTTGACTTTTCCCTATCTCCCTCAAGCAGTCTGGTCA 300
Db 381 GCTGTTCTCGAATTCCTCTTAAATTTGACTTTTCCCTATCTCCCTCAAGCAGTCTGGTCA 440
Qy 301 AGAGATTCCTAGTTCAATGCTTTATAAAGCTAATCGAATACTCAATCGTATCATCTC 360
Db 441 AGAGATTCCTAGTTCAATGCTTTATAAAGCTAATCGAATACTCAATCGTATCATCTC 500
Qy 361 TATCTCGGATTAACACTCTTTTGACGATACGAGGAGTTTCTTCTGCTGCTTCTAAG 420
Db 501 TATCTCGGATTAACACTCTTTTGACGATACGAGGAGTTTCTTCTGCTGCTTCTAAG 560
Qy 421 CTGAAGCGCGCAACTTGCACTGATTACATCTCTTTGCGTTTCAGACGATATCTCAAG 480
Db 561 CTGAAGCGCGCAACTTGCACTGATTACATCTCTTTGCGTTTCAGACGATATCTCAAG 620
Qy 481 AGAAGCAACGGCTATCTTGGGAGATGAGATCGAACTTCTTGGAGCAAAAATTCACGGTC 540
Db 621 AGAAGCAACGGCTATCTTGGGAGATGAGATCGAACTTCTTGGAGCAAAAATTCACGGTC 680
Qy 541 TTTGATGTTAGTACAGCCGAGCAGGAGATGCGAGAGCGGCTCTTCTAATTTTCATC 600
Db 681 TTTTATTTTATTCATACCTTATCATATATTCATATATTCATATCTCTCTTCTAATTTCAAC 740
Qy 601 AAAGTTTACCTAGAGTTCTCAGGGAAGTTACCCCATCGCTCACATTTTCATAC 654
Db 741 AAATCTTACCTATATATGCTCTCAATTAATGTTACCCCATCTCTCTCAATTCATAC 795

RESULT 7
BX837992
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
BX837992 597 bp mRNA linear EST 19-MAR-2001
BX837992 Arabidopsis developing seed Arabidopsis thaliana cdna
clone M39C9 5', mRNA sequence.
BX837992
BX837992.1 GI:9781588
EST.
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 597)
White, J.A., Todd, J., Newman, T., Focks, N., Girke, T., Martinez de
Ilarduya, O., Jaworski, J.G., Ohlrogge, J. and Benning, C.

TITLE A new set of Arabidopsis expressed sequence tags from developing seeds. The metabolic pathway from carbohydrates to seed oil

JOURNAL Plant Physiol. 124 (4), 1582-1594 (2000)

PUBMED 11115876

COMMENT Contact: Benning, C
Dept. of Biochemistry & Molecular Biology
Michigan State University
224 Biochemistry, Michigan State University, East Lansing, MI 48824, USA
Tel: 517 355 1609
Fax: 517 353 9334
Email: benning@msu.edu
Michigan State University DNA Sequencing Facility Arabidopsis Biological Resource Center, The Ohio State University, 309 Botany & Zoology Bldg., 1735 Neil Avenue, Columbus, OH 43210 USA, FAX: 6142920603 TEL: 6142929371.

FEATURES
source
1. .597
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/scotye="Columbia"
/db_xref="taxon:3702"
/clone="M39C9"
/tissue_type="seed"
/dev_stage="5-13 days after flowering"
/lab_host="E.coli"
/clone_lib="Arabidopsis developing seed"
/notes="Organ: Developing seed; Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoII"

ORIGIN
Query Match 52.0%; Score 592.8; DB 2; Length 597;
Best Local Similarity 99.5%; Pred. No. 1.8e-175;
Matches 594; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 519 CCTTGAACAAATTCACGGCTCTTTCATGCTAGTCTCAGCGGAGCGAGGATGCAGAA 578
DB 1 CCTTGAACAAATTCACGGCTCTTTCATGCTAGTCTCAGCGGAGCGAGGATGCAGAA 60

QY 579 GAGCCGCTCTTCTAAATTCATCAAGTTTCACCTAGAGTTCTTCAGGGAAGTTACCCCAT 638
DB 61 GAGCCGCTCTTCTAAATTCATCAAGTTTCACCTAGAGTTCTTCAGGGAAGTTACCCCAT 120

QY 639 CGCTCAATTCATAGCAGTTAAACGCTTTAGGCTCTCGGGGACCGAGGAAGATGCGTTG 698
DB 121 CGCTCAATTCATAGCAGTTAAACGCTTTAGGCTCTCGGGGACCGAGGAAGATGCGTTG 180

QY 699 CATCATGGATACATACCTATGAGCATCGTGAGTCCGCGAGGAGTAGTAGCTTCAACATC 758
DB 181 CATCATGGATACATACCTATGAGCATCGTGAGTCCGCGAGGAGTAGTAGCTTCAACATC 240

QY 759 CATAAGCTCTTTTCCAGTCCGTCATCCAGGCTCTTTAGGCTCTCACTCAAAACCATTCGG 818
DB 241 CATAAGCTCTTTTCCAGTCCGTCATCCAGGCTCTTTAGGCTCTCACTCAAAACCATTCGG 300

QY 819 CAGTAATAGTGCATCATGTAGCGACTCAGCAACAACTGGGAGATCCACCATTTGGTCT 878
DB 301 CAGTAATAGTGCATCATGTAGCGACTCAGCAACAACTGGGAGATCCACCATTTGGTCT 360

QY 879 GAGCAACAAGCTCCAGGTGGCATGAGCAGTTACGTTGCTGTGCTTAAATTTCCATGG 938
DB 361 GAGCAACAAGCTCCAGGTGGCATGAGCAGTTACGTTGCTGTGCTTAAATTTCCATGG 420

QY 939 TCAGGTCACAGTGGCTTCGGTTAAGAACTTTACGTTTGGCAGTTAGTGACGTGAAGC 998
DB 421 TCAGGTCACAGTGGCTTCGGTTAAGAACTTTACGTTTGGCAGTTAGTGACGTGAAGC 480

QY 999 AGGGCAGACATCTGAGAGGATCATATCTCCAGTTTGGGAAAGTTGGGAAGGACATGTTTAC 1058
DB 481 AGGGCAGACATCTGAGAGGATCATATCTCCAGTTTGGGAAAGTTGGGAAGGACATGTTTAC 540

QY 1059 CATGGATTATGGATATCCGATTTTCGGTTTCAAGCGTTTTCATCTGCTGAGCAG 1115
|||||

Db 541 CATGGATTATGGATATCCGATTTCTCGGTTTCAAGCGTTTGCTATCTGCTGAGCAG 597

RESULT 8
CNSOAI1DW 1281 bp mRNA linear HTC 06-FEB-2004
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone
DEFINITION GSLTFB702H02 of Flowers and buds of strain col-0 of Arabidopsis thaliana (thale cress).

ACCESSION BX830291
VERSION BX830291.1 GI:42458675
KEYWORDS HTC; GSLT cDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE
1. (bases 1 to 1281)
Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.
Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation
Unpublished
2. (bases 1 to 1281)
Genoscope.
Direct Submission
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

AUTHORS
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences) . 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full_length

FEATURES
source
1. .1281
/location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="GSLTFB702H02"
/tissue_type="Flowers and buds"
/scotye="Col-0"
/plasmid="pCWSFSPORT 6"
complement(1..1281)
/genes="At5g18680"

gene
complement(1..1281)
/genes="At5g18680"

ORIGIN
Query Match 49.9%; Score 569.2; DB 4; Length 1281;
Best Local Similarity 71.5%; Pred. No. 7.1e-168;
Matches 851; Conservative 0; Mismatches 268; Indels 71; Gaps 5;

QY 1 ATGACCTTCCGAAGTTTACTCCAGGAAATGCGGTCTTAGGCCACACCGGTGTAGTTACGCC 60
DB 11 ATGACCTTACGTAGCTTAATCTTGGAGTGGTTCGAGACCGCATCGTGTGTCCACGAC 70

QY 61 -----GCGCCTCAACCGCTAATAGTTTCAGACCCCTTCAGCTGCT 100
DB 71 CTGTGCGCGCGCGAGCTGCCGATTCCACTTCTGTGTATCCGAAGATTATCGTGTG 130

QY 101 CGGAGCTCCGGAGGAGCTGCTTAGAGAAATCTGTAGGTTAGAGCTGTGAGCTGTGACGCG 160
DB 131 CAGAGATTCTGAAGAGCTTCTTAGGAGATTCTGATTCTGTGTGAAGCGCGGACGCTG 190

161 GCGATTGGCCGTCGCGGCAAAACGTGGTCTGTGTCGGCGCTTTGTCTGTAGCTGGAGGA 220
Db |||||
191 GCGGATGGCCGTCACAGCCAGCGTGGTGGCTTTGTCGGCGCTTTGTGCTGGCGGC 250
|||
221 TTCTCACCAGGAGATTGAGTGTTCCTGGAATCTCTCTAAATCTCTCTAAATGACTTTCCCTATCT 280
Db |||||
251 TACTTATGAAGAAACCGTGGTGTCTCCCTGAGATCTCTCTAAATGACTTTCCCTCACTCT 310
|||
281 CCTCAGCAGTCTGGTCCAGAGATTCTCTAGTTCATGCTTTATTAACGATATCGAA 340
Db |||||
311 CTCTCAAGCAGCTGGTCCAGGGATTCTCTGGTTCATGCTTTATCAAAAGTATCGAA 370
|||
341 ATACTCAATCGTATCATCTCTATCTCGGATTAACCTCTTTGACGGATAACGGAGT 400
Db |||||
371 TTACGGCAATCATATCATCTCTATCTCGGATTAACCACTCTTTAACGGATGATGGAGT 430
|||
401 TTCTCTTCTGCTCTTAAGCTGMAAGCGCGCACTTGCACTGATTAATCATCTCTTTGTC 460
Db |||||
431 TTTTGTCTGCTGGTGAAGTTGAAGCACAACCTTGACGGATTACATTATCTCTTTAC 490
|||
461 GTTCAGCAGTATCTCAAGAGAGCAACGCGTATCTTGGGAGATGAGATCGAACTTCC 520
Db |||||
491 GTTCTGATGATGTGCGAAGAACCAAGCTTATGTTGGCAAGTGAGATCGAACTTCC 550
|||
521 TTGGACAAATTCACGGTCTTTGATGTAGTCAG-----ACCGGAGCAGCGAAGA 571
Db |||||
551 TAGGAACGAATTCATCTGCTTTTGTGGAATCTGCTGCTTCAACGGAGCGCGAAGT 610
|||
572 TGCAAGAGAGCGCTCTCTTAATTTTCAATCAAAAGTTTCACTAGAGTTCTCAGGGAAGTT 631
Db |||||
611 TGAGAAGAGCGCATCTTATATCCCGCAAAAGTTTCAGCAAAAGTTCTCTTTGGAAGTT 670
|||
632 ACCCCATCGCTCAATTCATAGAGATTAAAGTCTTAGGCTCTCGGAGCCGAGAGAA 691
Db |||||
671 ATCCTGCTCATATAACATATGAGTGAATGTCTTAGGATCCCGGGACCAAGAAGA 730
|||
692 TCGTTGTCATCATGATACAACTTACGAGATCGTGGAGTCGCGAGGAGTAGTCTT 751
Db |||||
731 TGCATATGCTTATGGACACAATACCTTACAGCAATATGGAGCTCAAGAGTAGTCTCAG 790
|||
752 CAACATCCATAAGCTCTTTTTCAGTCTGGTCTATCCAGTCTTTAGTCTCTACTCAAAAC 811
Db |||||
791 AACCATCAGAGTTTCCCTTACTCGGTACTCGGTCAACCTTATCCAGTCTCAGTCAAAAC 850
|||
812 CATTCGCGAGTATAGTGCATCATGAGCTCAGGCAACACCTGGGAGATCCACCAT 871
Db |||||
851 CATTCAGCAGTA-----GCTCAAGCCACCTGAAAGAAACACCAT 889
|||
872 TGGTGTGAGCAACAAGCTCCACGCTGGCATGAGCAGTTTACGTTCTGCTGCTTAAAT 931
Db |||||
890 TAGTGTGAGCAACAAGACACCGGTGGCAGCAGCAGTACCTGCTGCTGCTTGAAT 949
|||
932 TCATATGCTCAGTCAAGTGGCTTCGGTTAAGAACTTTTCAAGCTTTGTCGAG----- 982
Db |||||
950 TCATATGCTGTCACAGTAGCTCAGTGAAGAACTTTCAAGCTGCTGGCAGCAGGAGCTA 1009
|||
983 TTAGTGACTGTGAAGCAGGCGACATCTGAGA-----GGATCATCTCCAGT 1030
Db |||||
1010 GCTGTGCGAGTGGCAGCGGAATGTCAACGGAGAGCGAGCGAGCGGATTTATATTGCACT 1069
|||
1031 TTGGGAAGTTGGGAAGGACATCTTTACCATGATTTATGATATCCGATTTCTGCGTTTC 1090
Db |||||
1070 TTGGGAAGTGGCGGAAGATATGTTACGATGATTTATGATATACCGATCTCAGCTTTC 1129
|||
1091 AAGCGTTTGTCTGCTGAGCAGTTTGTGAACACCAAGATTTGCTGTGAA 1140
Db |||||
1130 AGCGTTTGCATTTGCTTGCAGCGCTCTGAGACTAGAACTCGTTGTGAA 1179
|||

RESULT 9
AV527654
LOCUS
DEFINITION AV527654 Arabidopsis thaliana aboveground organs two to six-week

old Arabidopsis thaliana cDNA clone AP242a03R 5', mRNA sequence.
AV527654 GI:8687182
EST.
AV527654.1
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 516)
Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
DNA Res. 7 (3), 175-180 (2000)
10907847
Contact: Brika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp. URL:http://www.kazusa.or.jp/en/plant/.
Location/Qualifiers
1..516
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Columbia"
/db_xref="taxon:3702"
/clone="AP242a03R"
/tissue_type="aboveground organs"
/dev_stage="two to six-week old"
/clone_lib="Arabidopsis thaliana aboveground organs two to
six-week old"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"
ORIGIN
Query Match 45.3%; Score 516; DB 1; Length 516;
Best Local Similarity 100.0%; Pred. No. 3.4e-151;
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 40 CCACACCGTGTAGTTTCAGCGCGCCCTCAACCGCTAATAGTTTCAGACCTTTACGCTGG 99
|||
Db 1 CCACACCGTGTAGTTTCAGCGCGCCCTCAACCGCTAATAGTTTCAGACCTTTACGCTGG 60
|||
QY 100 TCGGAGCTCCCGGAGGAGCTGCTTAGAGAAATCTCTGATTAGGGTTGAGACTGTGTACGCG 159
|||
Db 61 TCGGAGCTCCCGGAGGAGCTGCTTAGAGAAATCTCTGATTAGGGTTGAGACTGTGTACGCG 120
|||
QY 160 GCGGATTGGCCGCTCGCGCGAAACGTTGGTGGCTTTGCGCGGCTTTGTCGTAGCTGGAGG 219
|||
Db 121 GCGGATTGGCCGCTCGCGCGAAACGTTGGTGGCTTTGCGCGGCTTTGTCGTAGCTGGAGG 180
|||
QY 220 ATTCTCACAAGGAGATTGTAGCTGTTCTGAATCTCTCTAAATTTGACTTTCCCTATC 279
|||
Db 181 ATTCTCACAAGGAGATTGTAGCTGTTCTGAATTTCTCTCTAAATTTGACTTTCCCTATC 240
|||
QY 280 TCCCTCAGACGCTGTTCCAGAGATTCTCTAGTTCAATGCTTTTATAAAACGTAATCGA 339
|||
Db 241 TCCCTCAGACGCTGTTCCAGAGATTCTCTAGTTCAATGCTTTTATAAAACGTAATCGA 300
|||
QY 340 AATATCTCAATCTATCATCTCTATCTCGGATTAACCTACCTTTTACGGATAACCGGAAG 399
|||
Db 301 AATATCTCAATCTATCATCTCTATCTCGGATTAACCTACCTTTTACGGATAACCGGAAG 360
|||
QY 400 TTTCTTTTGTGCTTCTAAGCTGAAGCGGCAACTTGTGACCTGATTACATCATCTTTTG 459
|||
Db 361 TTTCTTTTGTGCTTCTAAGCTGAAGCGGCAACTTGTGACCTGATTACATCATCTTTTG 420
|||
QY 460 CGTTCAGACGATATCTCAAGAGAGACGCGGTATCTTTGGGAGATGAGATCGAACTTC 519
|||
Db 421 CGTTCAGACGATATCTCAAGAGAGACGCGGTATCTTTGGGAGATGAGATCGAACTTC 480
|||
QY 520 CTTGGAACAAAAATTACCGTCTTTTGATGTTAGTCAG 555
|||


```
QY 241 GCTGTTCTCGAATTCCTCTAAATGACTTTCCCTATCTCCCTCAAGCAGTCTGGTCCA 300
Db 387 GCTGTTCTCGAATTCCTCTAAATGACTTTCCCTATCTCCCTCAAGCAGTCTGGTCCA 446
QY 301 AGAGATTCTCTAGTTCATAGCTTTTATAAAACGTAATCGAAATCTCAATCGTATCATCTC 360
Db 447 AGAGATTCTCTAGTTCATAGCTTTTATAAAACGTAATCGAAATCTCAATCGTATCATCTC 506
QY 361 TATCTCGGATTAACTACCTCTTTTGACGGATAACGGGAAGTTTCTTCTTGCTGCTCTTAAG 420
Db 507 TATCTCGGATTAACTACCTCTTTTGACGGATAACGGGAAGTTTCTTCTTGCTGCTCTTAAG 566
QY 421 CTGAAGCGCGCAACTTGCACTGATTACATCAT 452
Db 567 CTGAAGCGCGCAACTTGCACTGATTACATCAT 598

RESULT 12
BU635249
LOCUS
DEFINITION
003805 Infected Arabidopsis Leaf Arabidopsis thaliana cDNA, mRNA
sequence.
ACCESSION
BU635249
VERSION
BU635249.1 GI:23302504
KEYWORDS
EST.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 704)
Lundsgaard,M., Emmersen,J., Nielsen,K.L., Wilson,I., Somerville,S.
and Welinder,K.G.
EST sequencing of Erysiphe cichoracearum infected Arabidopsis
Plants
Unpublished (2002)
Contact: Karen G. Welinder
Institut for bioteknologi
Aalborg Universitet
Sohngaardsholmsvej 49, 9000 Aalborg, Denmark
Tel: +45 96358467
Fax: +45 98141808
Email: kgw@bio.au.dk.
Location/Qualifiers
1. .704
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/dev_stage="Plant 3 weeks old, three days post infection"
/clone_lib="Infected Arabidopsis Leaf"
/notes="Organ: Leaf; Vector: pBluescript; Mixed cDNA
library of Arabidopsis and E. cichoracearum infected leaf
from three weeks old Arabidopsis plants. Plants were
harvested 3 days after infection and mRNA oligo dt
selected."

FEATURES
source
Query Match 37.1%; Score 422.6; DB 5; Length 704;
Best Local Similarity 76.8%; Pred. No. 1.3e-121;
Matches 534; Conservative 0; Mismatches 154; Indels 9; Gaps 1;

QY 58 GCCGCCGCTCAACCGCTAATAGTTCAGACCCCTTCAGCTGCTCGGAGCTCCCGAGGAG 117
Db 7 GCTGCCGATTCCTCTCTGTTGTCATCGCAAGATTATCGCTGGTTCAGAGATTCTCGAAGAG 66
QY 118 CTGCTTAGAGAAATCTTGATTAGGTTGAGACTGTTGACCGCGCGCAATGGCCGCTCGCG 177
Db 67 CTCTTTAGGAGATTCTGATTCTGTTGTTGAACGCGCGACGCTGGCGATGGCCGTCACGA 126
QY 178 CGAAACGTGCTGTGCGCGCTTTTGTGCTAGCTGGAGGATTCTCACCAAGGAGATT 237
```

```
Db 127 CGCAGCGTGGTGGCTGTGTCGCGGGGTTTGTCGTGCTGGCGGCTACTTATGAACGAACC 186
QY 238 GTAGCTGTTCTCGAATTCCTCTCTAAATTTGACTTTCCCTATCTCCCTCAAGCAGTCTGGT 297
Db 187 GTCGTTGTCCTCGAGATCTCTTTAAGTTGACTTTCCCATCTCTCTCAAGCAGCTGGT 246
QY 298 CCAAGAGATTCTCTAGTTCATAGTCTTATAAAACGTAATCGAAATCTCAATCGTATCAT 357
Db 247 CCAAGGATTCTCACTGGTTCAATGCTTTATCAACGTAATCGAATTTACGCAATCATCAT 306
QY 358 CTCATATCTCGAATTAACCTCTTTTGACGGATAACGGGAAGTTTCTTCTTGCTGCTTCT 417
Db 307 CTCTATCTCGGATTAAACCAACTCTTTTAACGGATGATGGGAAGTTTTCGTTGCTCGGCT 366
QY 418 AAGCTGAAGCGCGCAACTTGCACTGATTACATCATCTCTTTGCGTTTCAGACGATATCTCA 477
Db 367 AAGTTGAAGCACACAACTTGACGGATTACATTTATCTTTTACGTTCTGATGATATGCG 426
QY 478 AAGAGAAGCAACGGGTATCTTTGGGAGAAATGAGATCGAACTCTCTTGGAAACAAATTCAG 537
Db 427 AGAAGAAGCAAGCTTATGTTGGCAAGTGAGATCGAACTCTCTAGGAACGAAATTCCT 486
QY 538 GTCCTTTGATGTTAGTCTAG-----ACCGAGAGCAGCAAGATCGAAGAGCGCTCT 588
Db 487 GTCCTTTGATGTTAGTCTAGTCTCTTCAACGGGAGCGCAAGTTTGAGAAAGAGCGATCT 546
QY 589 TCTAATTTTCATCAAGTTTTCACCTAGAGTTCTCTCAGGGAAGTTTACCCATCGCTCACATT 648
Db 547 TATAATCCCGCAAAAGTTTCAGCAAAAGTCTCTTTGGAAGTTATCTCTGCTCTCATATC 606
QY 649 TCATACGAGTTTAAACGCTCTTAGGCTCTCGGGACCGAGAGAAGTTCGTTGCAATCATGGAT 708
Db 607 ACATATGAGCTCAATGCTTTAGGATCCCGGGGACCAAGAAAGATGCAATGCTTATGGAC 666
QY 709 ACAATPACTTATGAGCATCTGAGAGTCGAGAGTCGCGAGAGTAG 745
Db 667 ACAATPACTTACAGCAATGAGAGCTCAAGGAGTAG 703

RESULT 13
F13831
LOCUS
DEFINITION
ATT54753 Versaillies-VB Arabidopsis thaliana cDNA clone VBVC09 5',
similar to phosphodiesterase, mRNA sequence.
ACCESSION
F13831
VERSION
F13831.1 GI:757555
KEYWORDS
EST.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 420)
CNRS.
The Arabidopsis thaliana transcribed genome: the GDR cDNA program
Unpublished (1996)
Contact: Desprez T., Anselem J., Chiapello H., Rouze P., Caboche
M., Hoite H.
INRA Versailles
Laboratoire de Biologie Cellulaire
Route de Saint-Cyr, 78026 Versailles Cedex, France
Email: thierry@versaillies.inra.fr.
Location/Qualifiers
1. .420
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecot_type="Columbia"
/db_xref="taxon:3702"
/clone="VBVC09"
/tissue_type="whole seedlings"
/dev_stage="in vitro-grown etiolated seedlings, 5 days old"
/clone_lib="Versailles-VB"
/note="Vector: pBluescript"

FEATURES
source
```

[illegible]

COMMENT

On Jan 7, 1998 this sequence version replaced gi:948641.

Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.
Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22313tcn@bm.cl.msu.edu
Seq primer: T7.

FEATURES

source

Location/Qualifiers
1..358
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Columbia"
/db_xref="taxon:3702"
/clone="124E227"
/clone_lib="Lambda-PRL2"
/note="Vector: lambda Zip-Lox; Site 1: Sal; Site 2: Not;
Lambda PRL2 is a cDNA library derived from equal
quantities of 4 pools of mRNA. The mRNA sources were 1) 7
day germinated etiolated seedlings; 2) tissue culture
grown roots; 3) staged plants half with 24 hour light
cycle, half on 16 hr light, 8 hour dark- rosettes; 4)
same plants as 3 but aerial tissue (stems, flowers and
siliques. The vector is BRL's lambda Zip-Lox. The cDNA
inserts were directionally cloned with Sal-Not arms using
oligo dt primed cDNA. "

ORIGIN

Query Match 29.1%; Score 331.2; DB 8; Length 358;
Best Local Similarity 93.9%; Pred. No. 8.6e-93;
Matches 336; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 149 CTGTTGACGCGCGGATGGCCCTCGCGGCGAAACGTGGTGGCTTGTGCGCGGCTTTGTC 208
DB 1 CTGTTGACGCGCGGATGGCCCTCGCGGCGAAACGTGGTGGCTTGTGCGCGGCTTTGTC 60

QY 209 GTAGCTGGAGGATCTCACAGGAGATTGTAGCTGTTCTGTAATTCCTCTAAATTGA 268
DB 61 GTAGCTGGAGGATCTNACAGGAGATTGTAGCTGTTCTGTAATTCCTCTAAATTGA 120

QY 269 CTTTCCTATCTCCCTCAAGCACTGTGTCAGAGATTCTCTAGTTCAATGCTTTATAA 328
DB 121 CTTTCCTATCTCCCTCAAGCACTGTGTCAGAGATTCTNAGTTCAATGCTTTATAA 180

QY 329 AACGTAATCGAATACTCAATCGTATCATCTCTATCTCGGATTAACCTCTTTGACGG 388
DB 181 AACGTAATCGAATACTCAATCGTATCATCTCTATNNGGATTAACTCTTTNACGG 240

QY 389 ATAAACGGGAAGTTTCTTCTTGTCTTCTTAAGCTGAAGCGCGCAACTTGCACTGATTACA 448
DB 241 ATAAACGGGAAGTTTCTTCTTGTCTTCTTAAGCTGAAGCGCGCAACTNGCACTGATTACA 300

QY 449 TCATCTCTTTGCGTTTCAAGCATATCTCAAGAGAGCAACGCGTATCTTGGGAGAT 506
DB 301 TCATCTCTTTGCGTTTCAAGCGGTATCTCAAGGGGAGCAACGGTATCTTNGNGGAT 358

Search completed: December 26, 2005, 06:49:27
Job time : 5224 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 26, 2005, 03:25:36 ; Search time 267 Seconds
(without alignments)
7589.589 Million cell updates/sec

Title: US-10-763-042-20
Perfect score: 1140
Sequence: 1 atgacgtccgaagtact.....aaaccagaattgcctgtgaa 1140

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/1 COMB.seq.*
2: /cgn2_6/prodata/1/ina/5 COMB.seq.*
3: /cgn2_6/prodata/1/ina/6A COMB.seq.*
4: /cgn2_6/prodata/1/ina/6B COMB.seq.*
5: /cgn2_6/prodata/1/ina/H COMB.seq.*
6: /cgn2_6/prodata/1/ina/PCUS COMB.seq.*
7: /cgn2_6/prodata/1/ina/PP COMB.seq.*
8: /cgn2_6/prodata/1/ina/RE COMB.seq.*
9: /cgn2_6/prodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	72	6.3	299	3 US-09-313-294A-6623
2	66.8	5.9	1426	2 US-08-630-592-6
3	66.8	5.9	1426	2 US-08-714-991-6
4	66.8	5.9	1426	3 US-09-032-365A-7
5	66.8	5.9	1518	3 US-08-955-918C-11
6	66.8	5.9	1518	3 US-08-697-766A-11
7	66.8	5.9	1890	3 US-09-032-365A-61
8	66.8	5.9	1936	3 US-09-032-365A-59
9	66.8	5.9	2040	2 US-08-829-553-7
10	66.8	5.9	2040	2 US-08-922-267A-7
11	66.8	5.9	2040	2 US-08-936-707A-7
12	66.8	5.9	2040	2 US-08-936-706A-7
13	66.8	5.9	2040	3 US-09-248-203-7
14	66.8	5.9	2040	3 US-08-406-071-7
15	66.8	5.9	2040	3 US-08-955-918C-9
16	66.8	5.9	2040	3 US-08-697-766A-9
17	66.8	5.9	2040	3 US-09-814-986-7
18	66.8	5.9	2088	3 US-09-032-365A-64
19	66.8	5.9	2109	3 US-09-032-365A-63
20	66.8	5.9	2112	3 US-09-032-365A-56
21	66.8	5.9	2119	2 US-08-630-592-1
22	66.8	5.9	2119	2 US-08-714-991-1
23	66.8	5.9	2119	3 US-09-032-365A-1
24	66.8	5.9	2119	3 US-09-032-365A-1

25	66.8	5.9	2368	3 US-09-032-365A-57	Sequence 57, Appli
26	66.8	5.9	2434	2 US-08-630-592-3	Sequence 3, Appli
27	66.8	5.9	2434	2 US-08-714-991-3	Sequence 3, Appli
28	66.8	5.9	2434	3 US-09-032-365A-3	Sequence 3, Appli
29	66.8	5.9	3060	2 US-08-714-991-26	Sequence 26, Appli
30	66.8	5.9	3268	3 US-09-032-365A-9	Sequence 9, Appli
31	65.2	5.7	1512	3 US-08-955-918C-8	Sequence 8, Appli
32	65.2	5.7	1512	3 US-08-697-766A-8	Sequence 8, Appli
33	65.2	5.7	1801	3 US-08-955-918C-6	Sequence 6, Appli
34	65.2	5.7	1801	3 US-08-697-766A-6	Sequence 6, Appli
35	65.2	5.7	1804	2 US-08-631-200-1	Sequence 1, Appli
36	65.2	5.7	1804	2 US-08-829-553-1	Sequence 1, Appli
37	65.2	5.7	1804	2 US-08-922-267A-1	Sequence 1, Appli
38	65.2	5.7	1804	2 US-08-936-707A-1	Sequence 1, Appli
39	65.2	5.7	1804	2 US-08-936-706A-1	Sequence 1, Appli
40	65.2	5.7	1804	3 US-09-248-203-1	Sequence 1, Appli
41	65.2	5.7	1804	3 US-09-406-071-1	Sequence 1, Appli
42	65.2	5.7	1804	3 US-09-814-986-1	Sequence 1, Appli
43	65.2	5.7	2570	3 US-09-270-767-13498	Sequence 13498, A
44	64.6	5.7	1482	3 US-09-032-365A-16	Sequence 16, Appli
45	62.2	5.5	1743	3 US-09-032-365A-18	Sequence 18, Appli

ALIGNMENTS

RESULT 1
US-09-313-294A-6623
; Sequence 6623, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalugudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 6623
; LENGTH: 299
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700352120H1
; NAME/KEY: unsure
; LOCATION: 102, 134, 269
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-6623

Query Match 6.3%; Score 72; DB 3; Length 299;
Best Local Similarity 72.1%; Pred. No. 1.3e-13;
Matches 93; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
QY 1012 GAGAGATCATCTCCAGTTTCGGAAGTTGGGAAGACATGTTTACCATGATTGGA 1071
DB 2 GAAAGGTGATCCTGCAATTTGGGAAGTGTGGTGAAGACATGTTTACCATGACCT 61
QY 1072 TATCCGATTTTCGCTTTCAGCGTTTGTATCTGCTGAGCAGTTTTCGAACCAAGATT 1131
DB 62 TACCCACGTGTCAGCCTTCAGGCTTCGCGATCTGCTGANCAGCTTCGACCACGCTG 121
QY 1132 GCCTGTGAA 1140
DB 122 GCCTGTGAA 130

RESULT 2
US-08-630-592-6
; Sequence 6, Application US/08630592
; Patent No. 5770432

GENERAL INFORMATION:
; APPLICANT: Nishina, Patsy
; APPLICANT: No. 5770432enTrauth, Konrad
; APPLICANT: Naggert, Juergen
; APPLICANT: No. 5770432th, Michael
; TITLE OF INVENTION: Obesity Associated Genes
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
; STREET: 3400 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 941114187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MSDOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/630,592
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J.
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: A59504/BIR/PUS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 7811989
; TELEFAX: (415) 3983249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1426 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-630-592-6

Query Match 5.9%; Score 66.8; DB 2; Length 1426;
Best Local Similarity 57.1%; Pred. No. 1.9e-11;
Matches 153; Conservative 0; Mismatches 97; Indels 18; Gaps 1;

QY 872 TGGTCTGAGCAACAAAGCTCCACGGTGGCATGAGCAGTACGTTCTGGTGTAAATT 931
Db 1157 TCGAGCTGCAAAACAAGACACCTGTCTGGAATGATGACACACAGTCTATGTACTCAACT 1216

QY 932 TCCATGGTCGAGTCACAGTGGCTTCGGTTAAGAACTTTCAGCTTGTGGCAGTTAGTGACT 991
Db 1217 TCCATGGGCGGTCACACAGGCTCCGTGAAGAACTTCCAGATCATCCATGGCAATGACC 1276

QY 992 GTGAAGCAGGCGCAGACATCTGAGAGGATCATACTCCAGTTTGGGAAAGTTGGGAAGGACA 1051
Db 1277 -----CGGACTACATCGTGATGAGTTGGCGGTTAGCAGAGGATG 1318

QY 1052 TGTTTACCATGGATTATGGATATCCGATTTCTCGGTTTCAAGGTTTGTCTATCTGCCTGA 1111
Db 1319 TGTTCACCATGGATTACAACTACCCGCTGTGTGCACTGTCAGGCTTTGGCAATTGCCCTGT 1378

QY 1112 GCAGTTTGAACACAGAAATTCGCTGTGA 1139
Db 1379 CCAGCTTCGACAGCAAGCTGGCGTGCGA 1406

RESULT 3
US-08-714-991-6
; Sequence 6, Application US/08714991
; Patent No. 5776762
; GENERAL INFORMATION:
; APPLICANT: NISHINA, Michael
; APPLICANT: NISHINA, Patsy
; APPLICANT: No. 5776762en-Trauth, Konrad

; APPLICANT: NAGGERT, Juergen
; TITLE OF INVENTION: OBESITY ASSOCIATED GENES
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/714,991
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: SHERWOOD, Pamela J.
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: A-59504-1/PUS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-494-8700
; TELEFAX: 415-494-8771
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1426 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-714-991-6

Query Match 5.9%; Score 66.8; DB 2; Length 1426;
Best Local Similarity 57.1%; Pred. No. 1.9e-11;
Matches 153; Conservative 0; Mismatches 97; Indels 18; Gaps 1;

QY 872 TGGTCTGAGCAACAAAGCTCCACGGTGGCATGAGCAGTACGTTCTGGTGTAAATT 931
Db 1157 TCGAGCTGCAAAACAAGACACCTGTCTGGAATGATGACACACAGTCTATGTACTCAACT 1216

QY 932 TCCATGGTCGAGTCACAGTGGCTTCGGTTAAGAACTTTCAGCTTGTGGCAGTTAGTGACT 991
Db 1217 TCCATGGGCGGTCACACAGGCTCCGTGAAGAACTTCCAGATCATCCATGGCAATGACC 1276

QY 992 GTGAAGCAGGCGCAGACATCTGAGAGGATCATACTCCAGTTTGGGAAAGTTGGGAAGGACA 1051
Db 1277 -----CGGACTACATCGTGATGAGTTGGCGGTTAGCAGAGGATG 1318

QY 1052 TGTTTACCATGGATTATGGATATCCGATTTCTCGGTTTCAAGGTTTGTCTATCTGCCTGA 1111
Db 1319 TGTTCACCATGGATTACAACTACCCGCTGTGTGCACTGTCAGGCTTTGGCAATTGCCCTGT 1378

QY 1112 GCAGTTTGAACACAGAAATTCGCTGTGA 1139
Db 1379 CCAGCTTCGACAGCAAGCTGGCGTGCGA 1406

RESULT 4
US-09-032-365A-7
; Sequence 7, Application US/09032365A
; Patent No. 6114502
; GENERAL INFORMATION:
; APPLICANT: No. 6114502th, Michael
; APPLICANT: Nishina, Patsy
; APPLICANT: Naggert, Juergen
; APPLICANT: No. 6114502en-Trauth, Konrad
; TITLE OF INVENTION: GENE FAMILY ASSOCIATED WITH
; TITLE OF INVENTION: NEUROSENSORY DEFECTS
; NUMBER OF SEQUENCES: 67

```
;
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bozicevic & Reed, LLP
; STREET: 285 Hamilton Avenue, Suite 200
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/032,365A
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: SEQ-2CIP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-327-3400
; TELEFAX: 650 327-3231
; TELEX:
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1426 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; US-09-032-365A-7

Query Match 5.9%; Score 66.8; DB 3; Length 1426;
Best Local Similarity 57.1%; Pred. No. 1.9e-11;
Matches 153; Conservative 0; Mismatches 97; Indels 18; Gaps 1;

QY 872 TGGTGTGAGCAACAAAGCTCCACGGTGGCATGAGCAGTGTACGTTGCTGGTGTAAATT 931
DB 1157 TCGAGCTGCAAAACAGACACCTGCTGGATGATGACACACAGTCTATGTACTCACT 1216
QY 932 TCCATGGTCCAGTCCAGTGGCTTCGGTTAAGAACTTTCAGCTTTGGCAGTTAGTGACT 991
DB 1217 TCCATGGGCGCTCACACAGGCTCCGTGAAGAACTTCCAGATCATCCATGGCAATGACC 1276
QY 992 GTGAAGCAGGGCAGACATCTGAGAGGATCATCTCCAGTTTGGAAAGTTGGGAAGGACA 1051
DB 1277 -----CGGACTACATCGTGATGCAAGTTTGGCGGTAGCAGAGGATG 1318
QY 1052 TGTTCACATGGATTATGGATATCCGATTTCTCGGTTTCAAGGGTTTGCTATCTGCTGA 1111
DB 1319 TGTTCACATGGATTATCAACTACCCGCTGTGCACTGCACTGAGGCTTTGCCATTGCCCTGT 1378
QY 1112 GCAGTTTTGAAACCAAGATTGCTGTGA 1139
DB 1379 CCAGCTTCGACAGCAAGCTGGCGTGCGA 1406

RESULT 5
US-08-955-918C-11
; Sequence 11, Application US/08955918C
; Patent No. 6268130
; GENERAL INFORMATION:
; APPLICANT: Klymn, Patrick, and Moore, Karen
; TITLE OF INVENTION: RP Compositions and Therapeutic and
; NUMBER OF INVENTIONS: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts

;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bozicevic & Reed, LLP
; STREET: 285 Hamilton Avenue, Suite 200
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/955,918C
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/697,766
; FILING DATE: 29-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: MNI-007CPDV2CPA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1518 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1519
; US-08-955-918C-11

Query Match 5.9%; Score 66.8; DB 3; Length 1518;
Best Local Similarity 57.1%; Pred. No. 2e-11;
Matches 153; Conservative 0; Mismatches 97; Indels 18; Gaps 1;

QY 872 TGGTGTGAGCAACAAAGCTCCACGGTGGCATGAGCAGTGTACGTTGCTGGTGTAAATT 931
DB 1268 TCGAGCTGCAAAACAGACACCTGCTGGAATGATGACACACAGTCTATGTACTCACT 1327
QY 932 TCCATGGTCCAGTCCAGTGGCTTCGGTTAAGAACTTTCAGCTTTGGCAGTTAGTGACT 991
DB 1328 TCCATGGGCGCTCACACAGGCTCCGTGAAGAACTTCCAGATCATCCATGGCAATGACC 1387
QY 992 GTGAAGCAGGGCAGACATCTGAGAGGATCATCTCCAGTTTGGAAAGTTGGGAAGGACA 1051
DB 1388 -----CGGACTACATCGTGATGCAAGTTTGGCGGTAGCAGAGGATG 1429
QY 1052 TGTTCACATGGATTATGGATATCCGATTTCTCGGTTTCAAGGGTTTGCTATCTGCTGA 1111
DB 1430 TGTTCACATGGATTATCAACTACCCGCTGTGCACTGCACTGAGGCTTTGCCATTGCCCTGT 1489
QY 1112 GCAGTTTTGAAACCAAGATTGCTGTGA 1139
DB 1490 CCAGCTTCGACAGCAAGCTGGCGTGCGA 1517

RESULT 6
US-08-697-766A-11
; Sequence 11, Application US/08697766A
; Patent No. 6399760
; GENERAL INFORMATION:
; APPLICANT: Gimeno, Carlos; Klymn, Patrick; and Moore, Karen J.
; TITLE OF INVENTION: RP Compositions and Therapeutic and
; NUMBER OF INVENTIONS: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
```

;
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/697,766A
; FILING DATE: 29-AUG-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Silveri, Jean M.
; REGISTRATION NUMBER: 39,030
; REFERENCE/DOCKET NUMBER: MNI-007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1518 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1518
; US-08-697-766A-11

Query Match 5.9%; Score 66.8; DB 3; Length 1518;
Best Local Similarity 57.1%; Pred. No. 2e-11;
Matches 153; Conservative 0; Mismatches 97; Indels 18; Gaps 1;

QY 872 TGGTGTGAGCACAAGCTCCACGGTGGCATGACGAGTTAGCTTCTGGTGTAAAT 931
Db 1268 TCAGCTGCACAAACAGACACCTGCTGGAATGATGACACAGTCTTATGCTCACT 1327
QY 932 TCCATGTCGAGTCACAGTGGCTTCGGTTAAGAACTTTCAGCTTTGGCAGTTAGTGACT 991
Db 1328 TCCATGGCGCGTCACACAGGCTCCGTGAAGAACTTCCAGATCATCCATGCGCAATGACC 1387
QY 992 GTGAAGCAGGCGACAGATCTGAGAGATCATCTCCAGTTTGGAAAGTTGGAGGACA 1051
Db 1388 -----CGGACTACATCGTATGTCAGTTTGGCGGGTAGCAGAGGATG 1429
QY 1052 TGTTCACATGGATTATGGATATCCGATTCTCGCTTTCAAGCGTTTGTCTATCTGCTGA 1111
Db 1430 TGTTCACATGGATTACACTCCCGCTGTGTGCACTGCAGGCGCTTGCCATTTGCCCTGT 1489
QY 1112 GCAGTTTGAACACAGAAATTCCTGTGA 1139
Db 1490 CCAGCTTCGACAGCAAGCTGGCGTGCA 1517

RESULT 7
US-09-032-365A-61
; Sequence 61, Application US/09032365A
; Patent No. 6114502
; GENERAL INFORMATION:
; APPLICANT: No. 6114502th, Michael
; APPLICANT: Nishina, Patsy
; APPLICANT: Naggart, Juergen
; APPLICANT: No. 6114502en-Trauth, Konrad
; TITLE OF INVENTION: GENE FAMILY ASSOCIATED WITH
; TITLE OF INVENTION: GENE FAMILY ASSOCIATED WITH
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bozicevic & Reed, LLP
; STREET: 285 Hamilton Avenue, Suite 200
; CITY: Palo Alto

;
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/032,365A
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: SEQ-2CIP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-327-3400
; TELEFAX: 650 327-3231
; TELEX:
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1890 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-032-365A-61

Query Match 5.9%; Score 66.8; DB 3; Length 1890;
Best Local Similarity 57.1%; Pred. No. 2.2e-11;
Matches 153; Conservative 0; Mismatches 97; Indels 18; Gaps 1;

QY 872 TGGTGTGAGCACAAGCTCCACGGTGGCATGACGAGTTAGCTTCTGGTGTAAAT 931
Db 1271 TCGAGCTGCACAAACAGACACCTGCTGGAATGATGACACAGTCTTATGCTCACT 1330
QY 932 TCCATGTCGAGTCACAGTGGCTTCGGTTAAGAACTTTCAGCTTTGGCAGTTAGTGACT 991
Db 1331 TCCATGGCGCGTCACACAGGCTCCGTGAAGAACTTCCAGATCATCCATGCGCAATGACC 1390
QY 992 GTGAAGCAGGCGACAGATCTGAGAGATCATCTCCAGTTTGGAAAGTTGGAGGACA 1051
Db 1391 -----CGGACTACATCGTATGTCAGTTTGGCGGGTAGCAGAGGATG 1432
QY 1052 TGTTCACATGGATTATGGATATCCGATTCTCGCTTTCAAGCGTTTGTCTATCTGCTGA 1111
Db 1433 TGTTCACATGGATTACACTCCCGCTGTGTGCACTGCAGGCGCTTGCCATTTGCCCTGT 1492
QY 1112 GCAGTTTGAACACAGAAATTCCTGTGA 1139
Db 1493 CCAGCTTCGACAGCAAGCTGGCGTGCA 1520

RESULT 8
US-09-032-365A-59
; Sequence 59, Application US/09032365A
; Patent No. 6114502
; GENERAL INFORMATION:
; APPLICANT: No. 6114502th, Michael
; APPLICANT: Nishina, Patsy
; APPLICANT: Naggart, Juergen
; APPLICANT: No. 6114502en-Trauth, Konrad
; TITLE OF INVENTION: GENE FAMILY ASSOCIATED WITH
; TITLE OF INVENTION: GENE FAMILY ASSOCIATED WITH
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bozicevic & Reed, LLP
; STREET: 285 Hamilton Avenue, Suite 200
; CITY: Palo Alto

```
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FatSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/032,365A
; FILING DATE:
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: SEQ-2CIP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-327-3400
; TELEFAX: 650 327-3231
; TELEX:
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1936 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-032-365A-59

Query Match 5.9%; Score 66.8; DB 3; Length 1936;
Best Local Similarity 57.1%; Pred. No. 2.3e-11;
Matches 153; Conservative 0; Mismatches 97; Indels 18; Gaps 1;

QY 872 TGTGCTGACACAAAGCTCCACGGTGGCATGACGAGTTACCTTGTGCTTAAATT 931
DB 1316 TCGAGCTGCAAAACACACACCTGTCTGGAATGATGACACAGTCTCTATGTACTCAACT 1375
QY 932 TCCATGCTCAGTCACAGTGGCTTCGTTAAGAACTTTCAGCTTGTGGCAGTTAGTGACT 991
DB 1376 TCCATGGCGCGTCACACAGCCCTCGTGNAGAACTTCCAGATCATCCATGGCAATGACC 1435
QY 992 GTGAAGCAGGCGACACATCTGAGAGGATCATCTCCAGTTTGGAAAGTTGGGAAGACA 1051
DB 1436 -----CGGACTACATCGTGATGCGAGTTTGGCCGGTAGCAGAGGATG 1477
QY 1052 TGTTTACCATGATTATGGATATCCGATTTCTCGGTTTCAGCGTTTGCTATCTGCTGA 1111
DB 1478 TGTTCACCATGATTACAACTACCCGCTGTGTGCACTGCAGGCGTTTGCCATTGCCCTGT 1537
QY 1112 GCAGTTTGAACCCAGAAATTCGCTGTGA 1139
DB 1538 CCAGCTTCGACAGCAAGCTGGCGTGGA 1565

RESULT 9
US-08-631-200-7
; Sequence 7, Application US/08631200
; Patent No. 5646040
; GENERAL INFORMATION:
; APPLICANT: Kleyn, Patrick W.
; APPLICANT: Moore, Karen J.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.

; STATE: CA
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/631,200
; FILING DATE: 12-APR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-057
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2040 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 153..1670
; US-08-631-200-7

Query Match 5.9%; Score 66.8; DB 2; Length 2040;
Best Local Similarity 57.1%; Pred. No. 2.3e-11;
Matches 153; Conservative 0; Mismatches 97; Indels 18; Gaps 1;

QY 872 TGTGCTGACACAAAGCTCCACGGTGGCATGACGAGTTACCTTGTGCTTAAATT 931
DB 1420 TCGAGCTGCAAAACACACACCTGTCTGGAATGATGACACAGTCTCTATGTACTCAACT 1479
QY 932 TCCATGCTCAGTCACAGTGGCTTCGTTAAGAACTTTCAGCTTGTGGCAGTTAGTGACT 991
DB 1480 TCCATGGCGCGTCACACAGGCGCTCGTGAAGAACTTCCAGATCATCCATGGCAATGACC 1539
QY 992 GTGAAGCAGGCGACACATCTGAGAGGATCATCTCCAGTTTGGAAAGTTGGGAAGACA 1051
DB 1540 -----CGGACTACATCGTGATGCGAGTTTGGCCGGTAGCAGAGGATG 1581
QY 1052 TGTTTACCATGATTATGGATATCCGATTTCTCGGTTTCAAGCGTTTGCTATCTGCTGA 1111
DB 1582 TGTTCACCATGATTACAACTACCCGCTGTGTGCACTGCAGGCGTTTGCCATTGCCCTGT 1641
QY 1112 GCAGTTTGAACCCAGAAATTCGCTGTGA 1139
DB 1642 CCAGCTTCGACAGCAAGCTGGCGTGGA 1669

RESULT 10
US-08-829-553-7
; Sequence 7, Application US/08829553
; Patent No. 5817762
; GENERAL INFORMATION:
; APPLICANT: Kleyn, Patrick W.
; APPLICANT: Moore, Karen J.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
```

```
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/829,553
/ FILING DATE: 28-MAR-1997
/ CLASSIFICATION: 530
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/631,200
/ FILING DATE: 12-APR-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Coruzzi, Laura A.
/ REGISTRATION NUMBER: 30,742
/ REFERENCE/DOCKET NUMBER: 7853-057
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 790-9090
/ TELEFAX: (212) 869-9741/8864
/ TELEX: 66141 PENNIE
/ INFORMATION FOR SEQ ID NO: 7:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2040 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: unknown
/ TOPOLOGY: unknown
/ MOLECULE TYPE: DNA (genomic)
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 153..1670
/
US-08-829-553-7

Query Match 5.9%; Score 66.8; DB 2; Length 2040;
Best Local Similarity 57.1%; Pred. No. 2.3e-11;
Matches 153; Conservative 0; Mismatches 97; Indels 18; Gaps 1;

QY 872 TGGTGTGAGCAACAAAGCTCCACGGTGGCATGAGCAGTTAGCTTGCTGGTGTAAATT 931
Db 1420 TCGAGCTGCACAAACAGACACCTGCTGGAATGATGACACACAGTCTATGACTCAACT 1479
QY 932 TCATGGTTCGAGTACAGTGGCTTCGGTTAAGAACTTTTCAGCTTTGGGAGTTAGTACT 991
Db 1480 TCCATGGGCGCTGCACAGAGCTCCGTGAAGAACTTCCAGATCATCCATGGCAATGACC 1539
QY 992 GTGAAGCAGGGCAGACATCTGAGAGATCATCTCCAGTTTGGAAAGTTGGGAAGGACA 1051
Db 1540 -----CGGACTACATCGTATGATGAGTTGGCCGGTAGCAGAGGATG 1581
QY 1052 TGTTTACCATGGATTATGGATATCCGATTTCTGCGTTTCAAGCGTTTGCTATCTGCTGA 1111
Db 1582 TGTTCACATGGATTACAACTACCCGCTGTGTGCACTGCGAGGCTTTGCCATTGCCCTGT 1641
QY 1112 GCAGTTTGAACACAGAAATTCCTGTGA 1139
Db 1642 CCAGCTTCGACAGCAAGCTGGCGTGGA 1669

RESULT 11
US-08-922-267A-7
; Sequence 7, Application US/08922267A
; Patent No. 5861239
; GENERAL INFORMATION:
; APPLICANT: Klevn, Patrick W.
; APPLICANT: Moore, Karen J.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; NUMBER OF SEQUENCES: 82
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
;

/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/829,553
/ FILING DATE: 28-MAR-1997
/ CLASSIFICATION: 530
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/631,200
/ FILING DATE: 12-APR-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Coruzzi, Laura A.
/ REGISTRATION NUMBER: 30,742
/ REFERENCE/DOCKET NUMBER: 7853-057
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 790-9090
/ TELEFAX: (212) 869-9741/8864
/ TELEX: 66141 PENNIE
/ INFORMATION FOR SEQ ID NO: 7:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2040 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: unknown
/ TOPOLOGY: unknown
/ MOLECULE TYPE: DNA
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 153..1670
/
US-08-922-267A-7

Query Match 5.9%; Score 66.8; DB 2; Length 2040;
Best Local Similarity 57.1%; Pred. No. 2.3e-11;
Matches 153; Conservative 0; Mismatches 97; Indels 18; Gaps 1;

QY 872 TGGTGTGAGCAACAAAGCTCCACGGTGGCATGAGCAGTTAGCTTGCTGGTGTAAATT 931
Db 1420 TCGAGCTGCACAAACAGACACCTGCTGGAATGATGACACACAGTCTATGACTCAACT 1479
QY 932 TCATGGTTCGAGTACAGTGGCTTCGGTTAAGAACTTTTCAGCTTTGGGAGTTAGTACT 991
Db 1480 TCCATGGGCGCTGCACAGAGCTCCGTGAAGAACTTCCAGATCATCCATGGCAATGACC 1539
QY 992 GTGAAGCAGGGCAGACATCTGAGAGATCATCTCCAGTTTGGAAAGTTGGGAAGGACA 1051
Db 1540 -----CGGACTACATCGTATGATGAGTTGGCCGGTAGCAGAGGATG 1581
QY 1052 TGTTTACCATGGATTATGGATATCCGATTTCTGCGTTTCAAGCGTTTGCTATCTGCTGA 1111
Db 1582 TGTTCACATGGATTACAACTACCCGCTGTGTGCACTGCGAGGCTTTGCCATTGCCCTGT 1641
QY 1112 GCAGTTTGAACACAGAAATTCCTGTGA 1139
Db 1642 CCAGCTTCGACAGCAAGCTGGCGTGGA 1669

RESULT 12
US-08-936-707A-7
; Sequence 7, Application US/08936707A
; Patent No. 5871931
; GENERAL INFORMATION:
; APPLICANT: Klevn, Patrick W.
; APPLICANT: Moore, Karen J.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
```

ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/936,707A
FILING DATE: 24-SEP-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-100
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2040 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: CDS
LOCATION: 153..1670
US-08-936-707A-7

Query Match 5.9%; Score 66.8; DB 2; Length 2040;
Best Local Similarity 57.1%; Pred. No. 2.3e-11;
Matches 153; Conservative 0; Mismatches 97; Indels 18; Gaps 1;

QY 872 TGGTCTGAGCAACAAAGCTCCACGGTGGCATGAGCAGTTACGTTGCTGGTAAAT 931
Db 1420 TCGAGCTGCAAAACAAGACACCTGCTGGAATGATGACACACAGTCTATGACTCAACT 1479
QY 932 TCATGGTGCAGTCACAGTGGCTTCGGTTAAGAACTTTCAGCTTGGCAGTTAGTGACT 991
Db 1480 TCATGGGCGCTGACACAGGCTCGTGAAGAACTTCCAGATCATCCATGGCAATGACC 1539
QY 992 GTGAACGAGGCGACATCTGAGAGGATCATATCTCCAGTTTGGAAAGTTGGGAAGGACA 1051
Db 1540 -----CGGACTACATCGTATGATGAGTTTGGCGGGTAGCAGAGGATG 1581
QY 1052 TGTTTACCATGATATGATATCCGATTTTCGCTTCAAGCGTTTGCTATCTGCTGA 1111
Db 1582 TGTTCACCATGATATCAACTACCGCTGTGTGCACTGCGAGGCTTTGCCATTTGCCCTGT 1641
QY 1112 GCAGTTTTCGAACCAAGAAATTCGCTGGA 1139
Db 1642 CCAGCTTCGACAGCAAGCTGGCGTGCGA 1669

RESULT 13
US-08-936-706A-7
Sequence 7, Application US/08936706A
Patent No. 5876919
GENERAL INFORMATION:
APPLICANT: Kleyn, Patrick W.
APPLICANT: Moore, Karen J.
TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York

CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/936,706A
FILING DATE: 24-SEP-1997
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-099
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2040 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: CDS
LOCATION: 153..1670
US-08-936-706A-7

Query Match 5.9%; Score 66.8; DB 2; Length 2040;
Best Local Similarity 57.1%; Pred. No. 2.3e-11;
Matches 153; Conservative 0; Mismatches 97; Indels 18; Gaps 1;

QY 872 TGGTCTGAGCAACAAAGCTCCACGGTGGCATGAGCAGTTACGTTGCTGGTAAAT 931
Db 1420 TCGAGCTGCAAAACAAGACACCTGCTGGAATGATGACACACAGTCTATGACTCAACT 1479
QY 932 TCATGGTGCAGTCACAGTGGCTTCGGTTAAGAACTTTCAGCTTGGCAGTTAGTGACT 991
Db 1480 TCATGGGCGCTGACACAGGCTCGTGAAGAACTTTCAGATCATCCATGGCAATGACC 1539
QY 992 GTGAACGAGGCGACATCTGAGAGGATCATATCTCCAGTTTGGAAAGTTGGGAAGGACA 1051
Db 1540 -----CGGACTACATCGTATGATGAGTTTGGCGGGTAGCAGAGGATG 1581
QY 1052 TGTTTACCATGATATGATATCCGATTTTCGCTTCAAGCGTTTGCTATCTGCTGA 1111
Db 1582 TGTTCACCATGATATCAACTACCGCTGTGTGCACTGCGAGGCTTTGCCATTTGCCCTGT 1641
QY 1112 GCAGTTTTCGAACCAAGAAATTCGCTGGA 1139
Db 1642 CCAGCTTCGACAGCAAGCTGGCGTGCGA 1669

RESULT 14
US-09-248-203-7
Sequence 7, Application US/09248203
Patent No. 6043346
GENERAL INFORMATION:
APPLICANT: Kleyn, Patrick W.
APPLICANT: Moore, Karen J.
TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York

STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/248,203
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/936,707
FILING DATE: 24-SEP-1997
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-100
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2040 base pairs
TYPE: nucleic acid
STRADEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: CDS
LOCATION: 153..1670
US-09-248-203-7

Query Match 5.9%; Score 66.8; DB 3; Length 2040;

Best Local Similarity 57.1%; Pred. No. 2.3e-11;
Matches 153; Conservative 0; Mismatches 97; Indels 18; Gaps 1;
QY 872 TGGTGTGAGCAACAAAGCTCCACGGTGGCATGAGCAGGTACGTTGCTGGTCTTAAATT 931
Db 1420 TCGAGCTGCAAAACAAAGACACCTGTCTGGAATGATGACACACAGTCTTATGACTCAACT 1479
QY 932 TCCATGTCGAGTCACAGTGGCTTCGGTTAAGAACTTTCAGCTTGTGGCAGTTAGTGACT 991
Db 1480 TCCATGGGCGGTACACAGGCTCCGTGAAGAACTTCCAGATCATCCATGGCAATGACC 1539
QY 992 GTGAAGCAGGGCAGACATCTGAGAGGATCATATCTCCAGTTTGGGAAAGTTGGGAAGGACA 1051
Db 1540 -----CGGACTACATCGTGATGAGTTTGGCCGGGTAGCAGAGGATG 1581
QY 1052 TGTTCACCATGGATTATGATATCCGATTTTCGCGTTTCAAGCGTTTGTCTATCTGCCTGA 1111
Db 1582 TGTTCACCATGGATTATCAACTACCGCTGTGTGCACTGCGAGGCTTTGCCATTGCCCTGT 1641
QY 1112 GCAGTTTTCGAACCAAGATTGCTGTGA 1139
Db 1642 CCAGCTTCGACAGCAAGCTGGCGTGCGA 1669

RESULT 15

US-09-406-071-7
Sequence 7, Application US/09406071
Patent No. 6207386
GENERAL INFORMATION:
APPLICANT: Klevn, Patrick W.
APPLICANT: Moore, Karen J.
TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York

STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/406,071
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/936,707
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-100
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2040 base pairs
TYPE: nucleic acid
STRADEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: CDS
LOCATION: 153..1670
US-09-406-071-7

Query Match 5.9%; Score 66.8; DB 3; Length 2040;

Best Local Similarity 57.1%; Pred. No. 2.3e-11;
Matches 153; Conservative 0; Mismatches 97; Indels 18; Gaps 1;
QY 872 TGGTGTGAGCAACAAAGCTCCACGGTGGCATGAGCAGGTACGTTGCTGGTCTTAAATT 931
Db 1420 TCGAGCTGCAAAACAAAGACACCTGTCTGGAATGATGACACACAGTCTTATGACTCAACT 1479
QY 932 TCCATGTCGAGTCACAGTGGCTTCGGTTAAGAACTTTCAGCTTGTGGCAGTTAGTGACT 991
Db 1480 TCCATGGGCGGTACACAGGCTCCGTGAAGAACTTCCAGATCATCCATGGCAATGACC 1539
QY 992 GTGAAGCAGGGCAGACATCTGAGAGGATCATATCTCCAGTTTGGGAAAGTTGGGAAGGACA 1051
Db 1540 -----CGGACTACATCGTGATGAGTTTGGCCGGGTAGCAGAGGATG 1581
QY 1052 TGTTCACCATGGATTATGATATCCGATTTTCGCGTTTCAAGCGTTTGTCTATCTGCCTGA 1111
Db 1582 TGTTCACCATGGATTATCAACTACCGCTGTGTGCACTGCGAGGCTTTGCCATTGCCCTGT 1641
QY 1112 GCAGTTTTCGAACCAAGATTGCTGTGA 1139
Db 1642 CCAGCTTCGACAGCAAGCTGGCGTGCGA 1669

Search completed: December 26, 2005, 06:53:57

Job time : 269 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 26, 2005, 03:41:26 ; Search time 654 Seconds
(without alignments)
904.600 Million cell updates/sec

Title: US-10-763-042-20

Perfect score: 1140

Sequence: 1 atgagctccgaagtact.....aaaccagaattgcctgtgaa 1140

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4168288 seqs, 259477437 residues

Total number of hits satisfying chosen parameters: 8336576

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA, New.*

- 1: /cgn2_6/prodata/1/pubnpna/US08_NEW_PUB.seq.*
- 2: /cgn2_6/prodata/1/pubnpna/US06_NEW_PUB.seq.*
- 3: /cgn2_6/prodata/1/pubnpna/US07_NEW_PUB.seq.*
- 4: /cgn2_6/prodata/1/pubnpna/PCT_NEW_PUB.seq.*
- 5: /cgn2_6/prodata/1/pubnpna/US09_NEW_PUB.seq.*
- 6: /cgn2_6/prodata/1/pubnpna/US10_NEW_PUB.seq.*
- 7: /cgn2_6/prodata/1/pubnpna/US11_NEW_PUB.seq.*
- 8: /cgn2_6/prodata/1/pubnpna/US11_NEW_PUB.seq.*
- 9: /cgn2_6/prodata/1/pubnpna/US11_NEW_PUB.seq.*
- 10: /cgn2_6/prodata/1/pubnpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	53	4.6	924	6	US-10-750-185-31303
C 2	36	3.2	1293	6	US-10-525-710-7
C 3	35.6	3.1	100387	7	US-11-121-086-94
C 4	34.8	3.1	1161	6	US-10-667-295-179
C 5	33.8	3.0	1389	6	US-10-750-185-64773
C 6	33.8	3.0	110711	6	US-10-995-561-13264
C 7	33.6	2.9	315	6	US-10-802-796-127
C 8	33	2.9	2848	6	US-10-750-185-56373
C 9	32.6	2.9	867	6	US-10-750-185-26853
C 10	32	2.8	2195	6	US-10-750-185-41445
C 11	31.8	2.8	1172	6	US-10-750-185-55303
C 12	31.8	2.8	1904	6	US-10-750-185-57408
C 13	31.4	2.8	600	6	US-10-750-185-2925
C 14	31.4	2.8	783	6	US-10-750-185-57709
C 15	31.2	2.7	1587	6	US-10-750-185-25701
C 16	31.2	2.7	4608	6	US-10-821-234-137
C 17	31.2	2.7	151169	7	US-11-121-086-38
C 18	30.6	2.7	909	6	US-10-750-185-41557
C 19	30.6	2.7	1499	6	US-10-750-185-30575
C 20	30.6	2.7	3234	6	US-10-750-185-32891
C 21	30.6	2.7	195998	6	US-10-995-561-13489
C 22	30.6	2.7	212716	7	US-11-121-086-95
C 23	30.4	2.7	1015	6	US-10-750-185-39157

C	24	30.4	2.7	1536	6	US-10-750-185-50389	Sequence 50389, A
C	25	30.4	2.7	207600	7	US-11-112-908-31	Sequence 21, Appl
C	26	30	2.6	1849	6	US-10-750-185-28537	Sequence 28537, A
C	27	30	2.6	2594	6	US-10-750-185-63480	Sequence 63480, A
C	28	29.8	2.6	1207	6	US-10-750-185-40341	Sequence 40341, A
C	29	29.8	2.6	2967	6	US-10-750-185-33220	Sequence 33220, A
C	30	29.8	2.6	156544	7	US-11-121-086-81	Sequence 81, Appl
C	31	29.8	2.6	244196	6	US-10-995-561-13327	Sequence 13327, A
C	32	29.8	2.6	285300	6	US-10-857-780-6	Sequence 6, Appl
C	33	29.6	2.6	1482	7	US-11-137-463-13	Sequence 13, Appl
C	34	29.6	2.6	1620	6	US-10-750-185-53467	Sequence 53467, A
C	35	29.6	2.6	4104	6	US-10-131-826A-449	Sequence 449, App
C	36	29.4	2.6	2689	6	US-10-750-185-53445	Sequence 53445, A
C	37	29.4	2.6	157230	7	US-11-112-908-64	Sequence 64, Appl
C	38	29.4	2.6	170508	7	US-11-112-908-62	Sequence 62, Appl
C	39	29.4	2.6	173115	7	US-11-112-908-65	Sequence 65, Appl
C	40	29.4	2.6	182190	7	US-11-121-086-102	Sequence 102, App
C	41	29.4	2.6	185393	7	US-11-121-086-101	Sequence 101, App
C	42	29.4	2.6	186442	7	US-11-121-086-104	Sequence 104, App
C	43	29.2	2.6	378	6	US-10-793-626-1135	Sequence 1135, Ap
C	44	29.2	2.6	378	6	US-10-793-626-2657	Sequence 2657, Ap
C	45	29.2	2.6	931	6	US-10-750-185-30913	Sequence 30913, A

ALIGNMENTS

RESULT 1

US-10-750-185-31303/c
; Sequence 31303, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31303
; LENGTH: 924
; TYPE: DNA
; ORGANISM: Bovine 19866881111971
US-10-750-185-31303

Query Match	4.6%;	Score 53;	DB 6;	Length 924;
Best Local Similarity	62.4%;	Pred. No. 1.2e-07;		
Matches	83;	Conservative	0;	Mismatches 50;
Indels	0;	Gaps	0;	
QY	1007	CATCTGAGAGGATCATCTCCAGTTTGGGAAAGTTGGGAAGGACATGTTTACCATGGATT	1066	
Db	785	GAGCGGACTACTCTGCTCGATTCGATTTGGTCGTAGCCCCAGACACGTTCCACATGACT	726	
QY	1067	ATGGATATCCGATTTCGCTTTCAGCGTTTCAAGCGTTTGTATCTGCTGACGAGCAGTTTGAACCA	1126	
Db	725	TCGCGTTCCACTTTGCCCGCTCCAAGCCTTTGCCATCTGCTTGTCCAGTTTCGATGGGA	666	
QY	1127	GAATGCGCTGTGA	1139	
Db	665	AGCTGCGCATGTGA	653	

RESULT 2

US-10-525-710-7
; Sequence 7, Application US/10525710

Qy	971	AGCTTGGCAGTTAGTGCATCTGACAGCGGCGAGCATCTGAGAGGATCATATCTCCAGT	1039
Db	82657	GTCTTCTCTGGTGTCTTCTGACTTCTGCGTCCCAATATCTCTTCAAGTATCATTTCTTAAT	82598
Qy	1031	TTGGGAAAGTTGGGAAGGACATGTTTACCATGGATTATGATATCCGATTCTCGGTTTC	1090
Db	82597	TTGCTTAATTGACCAATATCTGTTTAAATCCCTTGTCTATATGTTTCTCTCTTTAT	82538
Qy	1091	AAGCGTTTGGCTATC 1104	
Db	82537	TAGCTTTCAGATC 82524	
RESULT 4			
US-10-667-295-179			
; Sequence 179, Application US/10667295			
; Publication No. US20050257293A1			
; GENERAL INFORMATION:			
; APPLICANT: Mascia, Peter			
; TITLE OF INVENTION: BIOLOGICAL CONTAINMENT SYSTEM			
; FILE REFERENCE: 11696-047001			
; CURRENT APPLICATION NUMBER: US/10/667,295			
; CURRENT FILING DATE: 2003-09-17			
; PRIOR APPLICATION NUMBER: US 60/411,823			
; PRIOR FILING DATE: 2002-09-17			
; NUMBER OF SEQ ID NOS: 263			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 179			
; LENGTH: 1161			
; TYPE: DNA			
; ORGANISM: Triticum aestivum			
; FEATURE:			
; NAME/KEY: misc feature			
; LOCATION: (1)-(1161)			
; OTHER INFORMATION: Ceres Seq. ID no. 12598265			
US-10-667-295-179			
Query Match 3.1%; Score 34.8; DB 6; Length 1161;			
Best Local Similarity 53.7%; Pred. No. 0.24;			
Matches 72; Conservative 0; Mismatches 62; Indels 0; Gaps 0;			
Qy	968	TTGAGCTTGTGGCAGTTAGTGCATCTGTAAGCAGGCGGAGATCTGAGAGGATCATATCTCC	1027
Db	926	TTGAGCGTGTGGCAAGGCTAACCGTTAAGATGGATGTGCATCTGAGGCTCTCATGCTTT	985
Qy	1028	AGTTTGGGAAGTTGGGAAGGACATGTTTACCATGGATTATGGATATCCGATTCTTCGCT	1087
Db	986	TGGCTGGAGATTTTGGACTGAATGTTTACAAGATATCTAGAACTGTTTAGCTATGC	1045
Qy	1088	TTCAGCGTTTGGCT 1101	
Db	1046	TTTAAGTTTGTAT 1059	
RESULT 5			
US-10-750-185-64773/c			
; Sequence 64773, Application US/10750185			
; Publication No. US20050260603A1			
; GENERAL INFORMATION:			
; APPLICANT: MMI GENOMICS, INC.			
; APPLICANT: Denise, Sue K.			
; APPLICANT: KERR, Richard			
; APPLICANT: ROSENFELD, David			
; APPLICANT: HOLM, Tom			
; APPLICANT: BATES, Stephen			
; APPLICANT: FANTIN, Dennis			
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS			
; FILE REFERENCE: MM1100-2			
; CURRENT APPLICATION NUMBER: US/10/750,185			
; CURRENT FILING DATE: 2003-12-31			
; PRIOR APPLICATION NUMBER: US 60/437,482			
; PRIOR FILING DATE: 2002-12-31			
; NUMBER OF SEQ ID NOS: 64922			

```
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 64773
; LENGTH: 1389
; TYPE: DNA
; ORGANISM: Bovine 1986680552958
US-10-750-185-64773

Query Match          3.0%; Score 33.8; DB 6; Length 1389;
Best Local Similarity 48.7%; Pred. No. 0.6;
Matches 92; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

QY 201 CGTTTCGTAGTCGAGGATTCACCAAGGAGATTGTAGCTGTTCCTCGAATTCCTC 260
Db 1046 CTTTGAAGAAGCTGATGCTAGTATAGAGAAATTGTCTGTAGATAATGCTCAA 987

QY 261 TAAATGTACTTCCCTATCTCCCTCAAGCAGTCTGTCGAAGAGATTCCTAGTCAATG 320
Db 986 GAAGATTACTGACTATTGCTCTCTGAGCAAAATATCAAGAATTTTGGTGACATGG 927

QY 321 CTTTATAAAGCTAATCGAATACTCAATCGTATCATCTCTATCTCGGATTAACTACCTC 380
Db 926 TTTAAGGAATAATCTTTAAATACTACCAGCATGCACACATCTTAGACATATGTTCCAG 867

QY 381 TTTGACGGA 389
Db 866 GTCCATGGA 858

RESULT 6
US-10-995-561-13264/c
; Sequence 13264, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13264
; LENGTH: 110711
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(110711)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-995-561-13264

Query Match          3.0%; Score 33.8; DB 6; Length 110711;
Best Local Similarity 54.4%; Pred. No. 10;
Matches 68; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 259 TCTAATAGCTTTCCCTATCTCCCTCAAGCAGTCTGGTCCAAGAGATTCCTAGTCAAA 318
Db 66003 TTTAAATTTGTTTCAGAGATTCCTCCCTACCTTTTAGTTCATAAAGATACTCTCCTTTAT 65944

QY 319 TCGTTTATAAAGCTAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 378
Db 65943 TGCCTTCTATTAACTAAGAGTTATCTTCCCTTACCTTACCTTACCTTACCTTACCTTAC 65884

QY 379 TCTTTT 383
Db 65883 CCTTT 65879

RESULT 7
US-10-802-796-127
; Sequence 127, Application US/10802796
; Publication No. US20050250104A1

; GENERAL INFORMATION:
; APPLICANT: COLE, STEWART
; APPLICANT: BUCHRIESEN-BROSCH, ROLAND
; APPLICANT: GORDON, STEPHEN
; APPLICANT: BILLAULT, ALAIN
; TITLE OF INVENTION: A METHOD FOR ISOLATING A POLYNUCLEOTIDE OF INTEREST
; TITLE OF INVENTION: FROM THE GENOME OF A MYCOBACTERIUM USING A BAC-BASED
; TITLE OF INVENTION: DNA LIBRARY. APPLICATION TO THE DETECTION OF
; TITLE OF INVENTION: MYCOBACTERIA.
; FILE REFERENCE: 05394.0011-00000
; CURRENT APPLICATION NUMBER: US/10/802,796
; CURRENT FILING DATE: 2004-03-18
; PRIOR APPLICATION NUMBER: US/09/673,476
; PRIOR FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: PCT/IB99/00740
; PRIOR FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: 09/060,756
; PRIOR FILING DATE: 1998-04-16
; NUMBER OF SEQ ID NOS: 743
; SOFTWARE: PatentIN Ver. 2.2
; SEQ ID NO 127
; LENGTH: 315
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (129)
; OTHER INFORMATION: a, t, c or g
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (153)
; OTHER INFORMATION: a, t, c or g
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (231)
; OTHER INFORMATION: a, t, c or g
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (258)
; OTHER INFORMATION: a, t, c or g
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (298)
; OTHER INFORMATION: a, t, c or g
US-10-802-796-127

Query Match          2.9%; Score 33.6; DB 6; Length 315;
Best Local Similarity 62.2%; Pred. No. 0.27;
Matches 51; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 105 GTCCCGGAGGAGCTCTTAGAGAAATCTAGAAATCTAGGTTAGACTGTTGACGGCGCGA 164
Db 214 GCGCGGAGGAGCTCTGCTNTGGATCCCGGCTGGCATTCGCGCTGTTGGCGGCGCG 273

QY 165 TTGGCGCTCGCGCGCAACGCTG 186
Db 274 GTGGTGGGGGGGCAACGCTG 295

RESULT 8
US-10-750-185-56373/c
; Sequence 56373, Application US/10750185
; Publication No. US200502603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
```

```

; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: NM11100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 56373
; LENGTH: 2848
; TYPE: DNA
; ORGANISM: Bovine 19866881140451
US-10-750-185-56373

Query Match 2.9% Score 33; DB 6; Length 2848;
Best Local Similarity 51.7%; Pred.No. 1.8; Indels 0; Gaps 0;
Matches 75; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 240 AGCTGTTCTCGAATTCCTCTCTAAATGACTTTCCCTATCTCCCTCAAGCAGTCTGGTCC 299
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2578 AGTGGACATGACTTAGCCTCTAAACAACAGTCATAGCTCTAGCTAAGACAGTAAGTTTG 2519

QY 300 AAGAGATTCTTAGTTCATAGCTTTATAAACGTAATCGAAATCTCAATCGTATCATCT 359
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2518 AAACAACCTTTTGTCTCTTTAGAAATTTAGAAAGAATTTTGTGTTCTTTTAGTGATTCATGA 2459

QY 360 CTATCTCGGATTAACTACCTCTTTG 384
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2458 CAATCTCCAAATACTAGTTTCTGG 2434

RESULT 9
US-10-750-185-26853/c
; Sequence 26853, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DeNise, Sue K.
; APPLICANT: KEER, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: NM11100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26853
; LENGTH: 867
; TYPE: DNA
; ORGANISM: Bovine 19866880569853
US-10-750-185-26853

```

```

US-10-750-185-41445
; Sequence 41445, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: PANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 41445
; LENGTH: 2195
; TYPE: DNA
; ORGANISM: Bovine 19866880956289
US-10-750-185-41445

Query Match          2.8%; Score 32; DB 6; Length 2195;
Best Local Similarity 50.0%; Pred. No. 3.3; Indels 0; Gaps 0;
Matches 80; Conservative 0; Mismatches 80;

QY      236  TTGTAGCTGTTCTCGAATTCTCCTAAAATTGACTTTCCCTATCTCCCTCAAGCAGTCTG 295
Db      1449  TTTTAGTTATCTCGTGTGTTTTTTTTTAATTTCTTATGAGTTACTCCAGCTACTAGTTTT 1508

QY      296  GTCCACAGAGATTCTCTAGTTCAATGCTTTTATAAAACGTAATCGAAATPACTCAATCGTATC 355
Db      1509  CTCACAGACAAGACAAATTTCTAGCTTTTCGAGAACAAATATGCAAAAGTATAAATAGTAGG 1568

QY      356  ATCTCTATCTCGGATTAACCTACCTCTTTTGACGGATAACGG 395
Db      1569  TTTATTTATAAAGAAGTGTGTGAAGTTTCATGAATTAATTCGTCG 1608

```

Db 850 TTTTCTTTAACTGTGAGGAAGTGCTATATAATATGACATCATYACCATTTCTTTCATAT 791
QY 369 ATTAACACTCTTTGACCGGATAACCGGAAGTTTCTTCTGCTGCTTCTTAAGCTGAAGCG 428
Db 790 TTTTTCACAAAACCTTTGGTGTAGCATATAATAAATTAATCTGACAGCTTCAATGCCAAAATG 731
QY 429 CGCAACTTGCACCTGATTAC 447
Db 730 TGAATTTTCTACTTAAAC 712

RESULT 12

US-10-750-185-57408/c
; Sequence 57408, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10750185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 57408
; LENGTH: 1904
; TYPE: DNA
; ORGANISM: Bovine 19866881175911
US-10-750-185-57408

Query Match 2.8%; Score 31.8; DB 6; Length 1904;
Best Local Similarity 48.6%; Pred. No. 3.6;
Matches 87; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
QY 205 TGTCTAGCTGGAGGATTTCTCACCAGGAGATTGTAGCTGTTCCTCTCAATCTCTCTTAA 264
Db 656 TGGACTAGCTGGTTCATTTCTCCTCTATCCGTTGTTATTGTAGCTTAGGCTCTAGGTTCA 597
QY 265 TTGACTTTCCCTATCTCCCTCAAGCAGTCTGGTCCAGAGATTCTCTAGTTCAATGCTTT 324
Db 596 AAACCATGTCAATCACACATTTGGAGCAGAGAGAGGATACCCCATCACAAAGTACA 537
QY 325 ATAAACGTAATCGAATACTCAATCGTATCATCTCTATCTCGGATTAACCTCTTT 383
Db 536 TTCAGCTGAGCAGCTTTTCAATACTTTTATATTTTCTTGAATTAATTAGAGCCTT 478

RESULT 13

US-10-750-185-2925/c
; Sequence 2925, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10750185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31

; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2925
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Bovine MMBT17358
US-10-750-185-2925

Query Match 2.8%; Score 31.4; DB 6; Length 600;
Best Local Similarity 48.0%; Pred. No. 2.3;
Matches 86; Conservative 1; Mismatches 92; Indels 0; Gaps 0;
QY 205 TGTCTAGCTGGAGGATTTCTCACCAGGAGATTGTAGCTGTTCCTCTCAATCTCTTAA 264
Db 393 TGGACTAGCTGGTTCATTTCTCCTCTATCCGTTGTTATTGTAGCTTAGGCTCTAGGTTCA 334
QY 265 TTGACTTTCCCTATCTCCCTCAAGCAGTCTGGTCCAGAGATTCTCTAGTTCAATGCTTT 324
Db 333 AAACCATGTCAATCACACATTTGGAGCAGAGAGGATACCCCATCACAAAGTACA 274
QY 325 ATAAACGTAATCGAATACTCAATCGTATCATCTCTATCTCGGATTAACCTCTTT 383
Db 273 TTCAGCTGAGCAGCTTTTCAATACTTTTATATTTTCTTGAATTAATTAGAGCCTT 215

RESULT 14

US-10-750-185-57709/c
; Sequence 57709, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10750185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 57709
; LENGTH: 783
; TYPE: DNA
; ORGANISM: Bovine 19866880388936
US-10-750-185-57709

Query Match 2.8%; Score 31.4; DB 6; Length 783;
Best Local Similarity 59.6%; Pred. No. 2.7;
Matches 53; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
QY 363 TCTCGATTAACTACTCTTTCACGAGTAACGGAGATTCTTCTGCTGCTTCTAAGCT 422
Db 719 TTTCCAGTTTCCCTCACCTTATTACCCATAATGCTGTTTCTCTCTGCTTGTCTGCT 660
QY 423 GAAGCGCGCAACTTGCACCTGATTACATCA 451
Db 659 GCAGAGTGGGAGCGGCACCTGTTTCAACA 631

RESULT 15

US-10-750-185-25701/c
; Sequence 25701, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David

```

; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: WM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 25701
; LENGTH: 1587
; TYPE: DNA
; ORGANISM: Bovine
US-10-750-185-25701

```

	Query Match	2.7%	Score 31.2;	DB 6;	Length 1587;
	Best Local Similarity	54.3%;	Pred. No. 5.1;		
	Matches	63;	Conservative	0;	Mismatches 53; Indels 0; Gaps 0;
Qy	341	ATACTCAATCGTATCATCTCTATCTCGGATTAACCTACTCTTTGACGGATAACGGGAGCT	400		
Db	663	ACACTCCATGTATCAGGATTCAAGGAAACATGACACCTCTGTCTCAAGAAAAACAAAAGG	604		
Qy	401	TTCTTCTTGGCTGCTTCTTAAGCTGAAGCGCGCACTTGCATGATTACATCATCTCT	456		
Db	603	ATCTTGTGACAGCATTTCAATTGAATTTCTCACCTGGCATCTCCACCATTAGCACT	548		

Search completed: December 26, 2005, 07:23:33
Job time : 657 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 25, 2005, 22:15:46 ; Search time 796 Seconds
(without alignments)
9544.912 Million cell updates/sec

Title: US-10-763-042-20
Perfect score: 1140
Sequence: 1 atgagctccgaagttaact.....aaaccagaattgcctgtgaa 1140

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq 21.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*
14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1140	100.0	1140	14	Adw64874 Arabidops
2	1140	100.0	1143	12	Adn73048 Thale cre
3	572.8	50.2	1419	12	Ado61994 Transcript
4	571.2	50.1	1295	3	Aac37041 Arabidops
5	567.8	49.8	1140	14	Adw64876 Arabidops
6	396	34.7	396	8	Abx62764 Arabidops
7	387.4	34.0	1681	14	Adw16693 Eucalyptu
8	367.8	32.3	1629	13	Adx13792 Plant ful
9	354.8	31.1	2586	14	Adw16694 Eucalyptu
10	342.6	30.1	1218	14	Adw64868 Arabidops
11	342.6	30.1	1221	12	Adn73404 Thale cre
12	336.2	29.5	1948	14	Aeb67645 Rice geno
13	336	28.6	469	6	Ab193451 Arabidops
14	324.6	28.5	1887	14	Adw17065 Pinus rad
15	322	28.2	1923	14	Adw17066 Pinus rad
16	321	28.2	2181	14	Adw17064 Pinus rad
17	282.8	24.8	2444	14	Adw17062 Pinus rad
18	254	22.3	1911	14	Adw18185 Pinus rad
19	254	22.3	1941	14	Adw18186 Pinus rad

20	250.4	22.0	1182	14	Adw64867 Arabidops
21	248.8	21.8	1501	14	Adw17063 Pinus rad
22	245.2	21.5	1991	14	Adw16695 Eucalyptu
23	229.2	20.1	1002	13	Adx36491 Plant ful
24	210.2	18.4	1164	14	Adw64871 Arabidops
25	195.4	17.1	1335	14	Adw64875 Arabidops
26	195.4	17.1	1338	10	Adx37190 Plant yie
27	195.4	17.1	1338	12	Adi44432 Plant tra
28	193.6	17.0	1792	14	Adw16699 Eucalyptu
29	193.6	17.0	1852	14	Adw16697 Eucalyptu
30	193	16.9	1137	14	Adw64872 Arabidops
31	192	16.8	1980	14	Adw16698 Eucalyptu
32	188.2	16.5	1380	3	Aac33976 Arabidops
33	184	16.1	1287	14	Adw64870 Arabidops
34	184	16.1	1567	12	Ado61990 Transcript
35	169.8	14.9	1910	13	Adx47926 Plant ful
36	169.8	14.9	1930	13	Adx48044 Plant ful
37	169.6	14.9	1822	14	Adw16696 Eucalyptu
38	164.8	14.5	1414	13	Adx54648 Plant ful
39	162	14.2	1926	13	Adx51972 Plant ful
40	160.8	14.1	1750	13	Adx59298 Plant ful
41	160.4	14.1	571	13	Acn49879 Cotton pr
42	160.4	14.1	1030	13	Adr60407 Cotton cd
43	159.2	14.0	1800	14	Aeb67826 Rice geno
44	148.4	13.0	1365	14	Adw64866 Arabidops
45	146.4	12.8	1599	13	Adx47638 Plant ful

ALIGNMENTS

RESULT 1
ADW64874
ID ADW64874 standard; cDNA; 1140 BP.

XX AC ADW64874;

XX DT 07-APR-2005 (first entry)

XX DE Arabidopsis thaliana TUBBY-like protein (TLP) 9 encoding cDNA.

XX KW Transgenic plant; salt tolerance; crop improvement; cold tolerance;
oxidative stress; drought resistance; TUBBY-like protein; plant;
XX KW chromosome 2; gene; ss.

XX OS Arabidopsis thaliana.

XX FH Key Location/Qualifiers

XX FT CDS 1..1140

XX FT /*tag= a

XX FT /product= "Arabidopsis thaliana TLP9 protein"

XX FT /partial

XX FT /note= "No stop codon"

XX US2005014266-A1.

XX PN 20-JAN-2005.

XX PD 21-JAN-2004; 2004US-00763042.

XX PR 21-JAN-2003; 2003US-0441380P.

XX PA (SINT-) ACAD SINICA.

XX PI Shaw J, Lai C;

XX XX WPI; 2005-080955/09.

XX DR P-PSDB; ADW64863.

XX DR GENBANK; AF487270.

XX PT New isolated Arabidopsis TUBBY-like proteins, useful for producing transgenic or transformed cells or plants having higher tolerance to salt, chilling, pathogens, oxidative stress, or water-deficit.

XX PS Claim 7; SEQ ID NO 20; 43pp; English.

CC The invention relates to Arabidopsis TUBBY-like proteins (TLP) and their

CC corresponding nucleic acid sequences. The invention also relates to a

CC transformed cell or a transgenic plant containing TLP nucleic acid and a

CC polypeptide and nucleic acids are useful for producing transgenic or

CC transformed cells or plants having higher tolerance to salt, chilling,

CC pathogens, oxidative stress or water-deficit. The present sequence is the

CC Arabidopsis thaliana TLP9 encoding cDNA. The Arabidopsis thaliana TLP9

CC gene is located at chromosome II.

XX SQ Sequence 1140 BP; 285 A; 265 C; 279 G; 311 T; 0 U; 0 Other;

Query Match 100.0%; Score 1140; DB 14; Length 1140;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACGTTCCGAAGTTTACTCCAGGAATGCGGTCTAGGCCACACCGTGTAGTTCACGCC 60

DB 1 ATGACGTTCCGAAGTTTACTCCAGGAATGCGGTCTAGGCCACACCGTGTAGTTCACGCC 60

QY 61 GCGCCTCAACCGCTAATAGTTTACAGACCTTTTTCAGCTGGTTCGGAGCTCCCGGAGGAGCTG 120

DB 61 GCGCCTCAACCGCTAATAGTTTACAGACCTTTTTCAGCTGGTTCGGAGCTCCCGGAGGAGCTG 120

QY 121 CTTAGAGAAATCTCGATTAGGGTTGAGACTGTTGACGGCGGAGTGGCGGTGCGGCGCA 180

DB 121 CTTAGAGAAATCTCGATTAGGGTTGAGACTGTTGACGGCGGAGTGGCGGTGCGGCGCA 180

QY 181 AACGTGGTGGTGTGCGCGGCTTTCGTAGCTGAGGATTTCTACCAAGGAGATTGTA 240

DB 181 AACGTGGTGGTGTGCGCGGCTTTCGTAGCTGAGGATTTCTACCAAGGAGATTGTA 240

QY 241 GCTGTTCTCGAATTCCTCTAATTTGACTTTCCTTATCCCTCAAGCAGTCTGTGCA 300

DB 241 GCTGTTCTCGAATTCCTCTAATTTGACTTTCCTTATCCCTCAAGCAGTCTGTGCA 300

QY 301 AGAGATTCTCTAGTTCAATGCTTTATAAAGCTAATCGAATACTCAATCGTATCATCTC 360

DB 301 AGAGATTCTCTAGTTCAATGCTTTATAAAGCTAATCGAATACTCAATCGTATCATCTC 360

QY 361 TATCTCGATTAACTACTCTTTGACGGATAACGGGAAGTTTCTTCTGCTGCTTCTAAG 420

DB 361 TATCTCGATTAACTACTCTTTGACGGATAACGGGAAGTTTCTTCTGCTGCTTCTAAG 420

QY 421 CTGAAGCGGCAACTTGCATGATTACATCATCTCTTTGGTTTCCAGCATATCTCAAAG 480

DB 421 CTGAAGCGGCAACTTGCATGATTACATCATCTCTTTGGTTTCCAGCATATCTCAAAG 480

QY 481 AGAAGCAACGCGTATCTTGGGAGAAATGAGATCGAACTTCTTGGAAACAAATTCACGGTC 540

DB 481 AGAAGCAACGCGTATCTTGGGAGAAATGAGATCGAACTTCTTGGAAACAAATTCACGGTC 540

QY 541 TTTGATGTAGTCAGACCGGAGAGGAGATCGAGATCGAGACGCGCTCTCTAATTTTCATC 600

DB 541 TTTGATGTAGTCAGACCGGAGAGGAGATCGAGATCGAGACGCGCTCTCTAATTTTCATC 600

QY 601 AAAGTTTCACTAGATTCTCTCAGGGAAGTTTACCCTATGCTCACAATTTTATACAGATTGA 660

DB 601 AAAGTTTCACTAGATTCTCTCAGGGAAGTTTACCCTATGCTCACAATTTTATACAGATTGA 660

QY 661 AACGTCTTAGGCTCTCGGGACCGGAGAAATGCGTTTGCATCATGATCAATACCTATG 720

DB 661 AACGTCTTAGGCTCTCGGGACCGGAGAAATGCGTTTGCATCATGATCAATACCTATG 720

QY 721 AGCATGTGAGTGTGCGAGGAGTAGTGTTCACATCCGTAAGCTCTTTTCCAGTCGG 780

DB 721 AGCATGTGAGTGTGCGAGGAGTAGTGTTCACATCCGTAAGCTCTTTTCCAGTCGG 780

QY 781 TCATCACCAGTCTTAGTCTCACTCAAAACCAATGCGCAGTAAATAGTGCATCATGTAGC 840

DB 781 TCATCACCAGTCTTAGTCTCACTCAAAACCAATGCGCAGTAAATAGTGCATCATGTAGC 840

DB 781 TCATCACCAGTCTTAGTCTCACTCAAAACCAATGCGCAGTAAATAGTGCATCATGTAGC 840

QY 841 GACTCAGGCAACAACCTCGGAGATCCACATTTGGTGTGAGCAACAAGCTCCACGGTGG 900

DB 841 GACTCAGGCAACAACCTCGGAGATCCACATTTGGTGTGAGCAACAAGCTCCACGGTGG 900

QY 901 CATGAGCAGTTACGTTGCTGGTCTTAAATTTCCATGGTTCGAGTCCAGTGGCTTTCGGTT 960

DB 901 CATGAGCAGTTACGTTGCTGGTCTTAAATTTCCATGGTTCGAGTCCAGTGGCTTTCGGTT 960

QY 961 AAGAACTTTCAAGCTTTGCTGGTGTAGTGCATGCTGTAAGCAGGCGCAGATCTGAGAGATC 1020

DB 961 AAGAACTTTCAAGCTTTGCTGGTGTAGTGCATGCTGTAAGCAGGCGCAGATCTGAGAGATC 1020

QY 1021 ATACTCCAGTTTGGGAAAGTTGGGAGGACATGTTTACCATGATTTATGATATCCGATT 1080

DB 1021 ATACTCCAGTTTGGGAAAGTTGGGAGGACATGTTTACCATGATTTATGATATCCGATT 1080

QY 1081 TCTGCGTTTCAAGCGTTTGTCTATCTGCTGAGCAGTTTTGAACCAAGATTCGCTGTGAA 1140

DB 1081 TCTGCGTTTCAAGCGTTTGTCTATCTGCTGAGCAGTTTTGAACCAAGATTCGCTGTGAA 1140

RESULT 2

ADN73048

ID ADN73048 standard; cDNA; 1143 BP.

XX AC ADN73048;

XX XX

DT 15-JUL-2004 (first entry)

XX Thale cress cDNA upregulated in E2Fa/Dpa expressing plants SeqID 943.

DE gene; ss; plant; transgenic; E2Fa/Dpa transcription factor;

XX growth regulator; animal feed product; thale cress;

KW Cell wall biosynthesis; nitrogen metabolism; carbon metabolism.

KW Arabidopsis thaliana.

OS WO2004035798-A2.

XX XX

PN 29-APR-2004.

XX XX

PD 20-OCT-2003; 2003WO-EP011658.

XX XX

PP 18-OCT-2002; 2002EP-00079408.

XX XX

PR (CROP-) CROPDESIGN NV.

XX Inze D, De Veylder L, Vlieghe K;

XX WPI; 2004-348466/32.

DR P-PSDB; ADN73049.

XX XX

PT Altering plant characteristics, useful for producing plants for enzyme or

PT pharmaceutical production comprises modifying in a plant, expression of

PT one or more nucleic acids and/or modifying level or activity of one or

PT more proteins.

XX Claim 1; SEQ ID NO 943; 134pp; English.

PS This invention relates to a novel method for altering one or more plant

XX characteristics. Specifically, it refers to identifying genes that are up

CC - or down-regulated in transgenic plants overexpressing the heterodimeric

CC E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to

CC alter plant characteristics accordingly. The present invention describes

CC generating transgenic plants for the production of growth regulators,

CC enzymes, therapeutics, pharmaceuticals and animal feed products, where

CC the altered plant characteristics are selected from increased yield or

CC biomass, enhanced survival capacity, stress tolerance, plant architecture

CC or physiology, altered endoreplication, biochemistry, signal

CC transduction, storage lipid mobilisation and/or altered photosynthesis,

CC each relative to the corresponding wild type plants. Accordingly, these

CC sequences can also be useful as positive or negative selectable markers
CC during transformation of cells or tissues. The identified genes play a
CC role in a variety of biological processes such as DNA replication, cell
CC wall biosynthesis, nitrogen and/or carbon metabolism or they function as
CC transcription factors. This polynucleotide sequence is the cDNA
CC upregulated 1.3 fold or more in plants overexpressing the E2Fa/Dpa
CC transcription factor, given in an exemplification of the invention.
XX
SQ Sequence 1143 BP; 287 A; 265 C; 279 G; 312 T; 0 U; 0 Other;

Query Match 100.0%; Score 1140; DB 12; Length 1143;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACGTTCCGAGATTACTCCAGGAAATGCGGTCTAGGCCAACCGTGTAGTTTCAGCC 60
DB 1 ATGACGTTCCGAGATTACTCCAGGAAATGCGGTCTAGGCCAACCGTGTAGTTTCAGCC 60

QY 61 GCGGCTCAACCGCTAATAGTTTACAGACCTTTTCCAGCTGGTCGAGCTCCGGAGGAGCTG 120
DB 61 GCGGCTCAACCGCTAATAGTTTACAGACCTTTTCCAGCTGGTCGAGCTCCGGAGGAGCTG 120

QY 121 CTTAGAGAAATCTGATTAGGCTTGAGACTGTTGACGGCGGCGATTGGCGCTCGCGCGCA 180
DB 121 CTTAGAGAAATCTGATTAGGCTTGAGACTGTTGACGGCGGCGATTGGCGCTCGCGCGCA 180

QY 181 AACGTGGTGGTGTGTCGCGCGCTTGTGAGTGGAGATTCTCAACGAGGAGATTGTA 240
DB 181 AACGTGGTGGTGTGTCGCGCGCTTGTGAGTGGAGATTCTCAACGAGGAGATTGTA 240

QY 241 GCTGTCTCTGAATCTCTCTAAATGACTTCCCTATCTCCCTCAAGCAGTCTGTGCA 300
DB 241 GCTGTCTCTGAATCTCTCTAAATGACTTCCCTATCTCCCTCAAGCAGTCTGTGCA 300

QY 301 AGAGATTCTCTAGTTCAATGCTTTATAAAGCTTAATCGAAATCTCAATCGTATCATCTC 360
DB 301 AGAGATTCTCTAGTTCAATGCTTTATAAAGCTTAATCGAAATCTCAATCGTATCATCTC 360

QY 361 TATCTCGAATTAATACCTCTTTTGACGAGTAACGGGAAGTTCTTCTGCTGCTCTAAG 420
DB 361 TATCTCGAATTAATACCTCTTTTGACGAGTAACGGGAAGTTCTTCTGCTGCTCTAAG 420

QY 421 CTGAGCGCGCACTTGCACTGATTAATCATCTCTTTGCGTTCAGACGATATCTCAAG 480
DB 421 CTGAGCGCGCACTTGCACTGATTAATCATCTCTTTGCGTTCAGACGATATCTCAAG 480

QY 481 AGAAGCAACCGGTATCTCGGAGATGAGATCGAACTTCTTGGAACAAATTCACGGTC 540
DB 481 AGAAGCAACCGGTATCTCGGAGATGAGATCGAACTTCTTGGAACAAATTCACGGTC 540

QY 541 TTTGATGGTGTAGTCAGACCGGAGCAGGAAAGATGACAGAGCGGCTCTTCTAATTTTATC 600
DB 541 TTTGATGGTGTAGTCAGACCGGAGCAGGAAAGATGACAGAGCGGCTCTTCTAATTTTATC 600

QY 601 AAAGTTTACCTAGATTCTTACGGAAGTTTACCCATGCTCAATTTTACAGATTGTA 660
DB 601 AAAGTTTACCTAGATTCTTACGGAAGTTTACCCATGCTCAATTTTACAGATTGTA 660

QY 661 AACGCTTAGGCTCTCGGAGCGGAGAAAGATGCGTTGATCATGATACAAATCTATG 720
DB 661 AACGCTTAGGCTCTCGGAGCGGAGAAAGATGCGTTGATCATGATACAAATCTATG 720

QY 721 AGCATCGTGGAGTCCGAGGAGTAGTGTCTTCAACATCCATAGCTCTTTTCCAGTCGG 780
DB 721 AGCATCGTGGAGTCCGAGGAGTAGTGTCTTCAACATCCATAGCTCTTTTCCAGTCGG 780

QY 781 TCATCACCAAGTCTTTAGGTCTCACTCAAAACCATTCGGCAGTAATAGTGCATCATAGC 840
DB 781 TCATCACCAAGTCTTTAGGTCTCACTCAAAACCATTCGGCAGTAATAGTGCATCATAGC 840

QY 841 GACTCAGGCAACACTGGAGATCCACATTTGGTCTGAGCAACAAAGCTCCAGCTGG 900
DB 841 GACTCAGGCAACACTGGAGATCCACATTTGGTCTGAGCAACAAAGCTCCAGCTGG 900

QY 901 CATGACAGTTAGCTTGTGCTGCTTAAATTTCCATGGTCGAGTCACAGTGGCTTGGTT 960
DB 901 CATGACAGTTAGCTTGTGCTGCTTAAATTTCCATGGTCGAGTCACAGTGGCTTGGTT 960

QY 961 AAGAACTTTTCAAGCTTGTGCGAGTTAGTACTGTGAAGCAGGGCAGACATCTGAGAGGATC 1020
DB 961 AAGAACTTTTCAAGCTTGTGCGAGTTAGTACTGTGAAGCAGGGCAGACATCTGAGAGGATC 1020

QY 1021 ATACTCCAGTTTGGGAAAGTTGGGAAGGACATGTTTACCATGATTTATGGATATCCGATT 1080
DB 1021 ATACTCCAGTTTGGGAAAGTTGGGAAGGACATGTTTACCATGATTTATGGATATCCGATT 1080

QY 1081 TCTGCGTTTCAAGCGTTTCTCTATCTGCTGAGCAGTGTGAAACCAAGATTTGCTGTGAA 1140
DB 1081 TCTGCGTTTCAAGCGTTTCTCTATCTGCTGAGCAGTGTGAAACCAAGATTTGCTGTGAA 1140

RESULT 3
ADO61994
ID ADO61994 standard; DNA; 1419 BP.
XX
AC ADO61994;
XX
DT 15-JUL-2004 (first entry)
XX
DE Transcription factor G2603 coding sequence, SEQ ID 461.
XX
KW Plant; transcription factor; transgenic plant; abiotic stress tolerance;
KW osmotic stress tolerance; cold tolerance; heat tolerance;
KW low nitrogen tolerance; low phosphate tolerance; fungal disease;
KW glyphosate resistance; flowering; fertility; seed development; ds.
XX
OS Arabidopsis thaliana.
XX
PN WO2004031349-A2.
XX
PD 15-APR-2004.
XX
PF 18-SEP-2003; 2003WO-US030292.
XX
PR 18-SEP-2002; 2002US-0411837P.
PR 17-DEC-2002; 2002US-0434166P.
PR 24-APR-2003; 2003US-0465809P.
XX
PA (MEND-) MENDEL BIOTECHNOLOGY INC.
XX
XX Jiang C. Heard JE, Ratcliffe O, Creelman RA, Adam LJ, Reuber TL;
PI Riechmann JL, Haake V, Dubell AN, Keddie JS, Sherman BK;
XX
DR WPI; 2004-330163/30.
DR P-PSDB; ADO61995.
XX
PT New recombinant polynucleotide encoding transcription factor
PT polypeptides, useful for producing transgenic plants with advantageous
PT properties compared to a reference plant.
XX
PS Claim 1; SEQ ID NO 461; 510pp; English.
XX
CC The present invention relates to novel plant transcription factor
CC proteins (I) and nucleotide sequences (II) (ADO61534-ADO63778). The
CC sequences can be used to produce transgenic plants, which overexpress
CC (II), where the transgenic plant has an altered trait as compared to a
CC non-transgenic plant or wild-type plant. The transgenic plant comprises
CC an altered trait selected from increased tolerance to abiotic stress,
CC increased tolerance to osmotic stress, increased tolerance to cold,
CC increased germination in cold, increased tolerance to heat, increased
CC germination in heat, increased tolerance to freezing conditions,
CC increased tolerance to low nitrogen conditions, increased tolerance to
CC low phosphate conditions, increased tolerance to disease, including
CC fungal disease and particularly Erysiphe, Fusarium and Botrytis,
CC increased tolerance to multiple fungal pathogens, increased resistance to
CC glyphosate, increased sensitivity to ABA, reduced sensitivity to ABA,

CC increased sensitivity to ACC, altered sugar sensing, increased tolerance
CC to sugars, altered carbon/nitrogen sensing, early flowering, late
CC flowering, altered flower structure, loss of flower determinacy, reduced
CC fertility, altered shoot meristem development, altered branching pattern,
CC altered stem morphology, altered vascular tissue structure, reduced
CC apical dominance, altered trichome density, altered trichome development,
CC altered trichome structure, altered root development, altered shade
CC avoidance, altered seed development, altered seed ripening, altered seed
CC germination, slow growth, fast growth, altered cell differentiation,
CC altered cell proliferation, altered cell expansion, altered phase change,
CC altered senescence, abnormal embryo development, altered programmed cell
CC death, lethality when overexpressed, altered necrosis patterns, increased
CC plant size, increased biomass, large seedlings, dwarfed plants, dark
CC green leaves, change in leaf shape, increased leaf size and mass, light
CC green or gray leaves, glossy leaves, altered abaxial/adaxial polarity,
CC altered seed coloration, altered seed size, altered seed shape, large
CC seed, increased leaf wax, increased leaf fatty acids, altered seed oil
CC content, altered seed protein content, altered seedprenyl content,
CC altered leaf prenyl lipid content, increased anthocyanin levels, and
CC decreased anthocyanin levels. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1419 BP; 361 A; 317 C; 331 G; 410 T; 0 U; 0 Other;

Query Match 50.2%; Score 572.8; DB 12; Length 1419;
Best Local Similarity 71.7%; Pred. No. 1.6e-177;
Matches 852; Conservative 0; Mismatches 267; Indels 69; Gaps 5;

QY 1 ATGACGTCGAGTTTACTCCAGGAATGCGGTCTAGGCGCACCGTGTAGTTCACGCG 60
DB 66 ATGACCTTACGTAGCTTAATCTCTGAGATGCGTTCGAGACCGCATCGTGTGTCACGAC 125
QY 61 -----GCCGCCCTCAACCGCTAATAGTTTCAGACCCCTTTCAGCTGTGTCG 102
DB 126 CTTGCCGCCGCCGACGTCGCCGATTCACCTTCTGTGTCATCGCAAGATTATCGTGTCA 185
QY 103 GAGCTCCCGGAGAGCTGCTTTAGAGAAATCCTGTATTAGGGTTGAGACTGTTGACCGCGGC 162
DB 186 GAGATTCTCTGAAGAGCTTCTTAGGAGATTCGTATTCGTGTGTTGAAGCGCGGACGCTGCG 245
QY 163 GATTGCCCTCGCGGAAACGCTGCTGCTGCGCGGCTTGTGCTAGCTGGAGATT 222
DB 246 GGATGCCCTCGACGACGCGGTGGCTTGTGCGCGGCTTGTGCTGGCTGGCGGCTA 305
QY 223 CTCACCAAGGAGATTGTAGCTGTTCTGAATTCCTCTAAATTGACTTTCCTATCTCC 282
DB 306 CTTATGAAAGAACCTGCTGTTGCTCGTGAATCTCTTAAGTTGACTTTCCTCATCTCT 365
QY 283 CTCACGAGTCTGGTCCAGAGATTCTCTAGTTCAATGCTTTTATAAAAGTAAATCGAAAT 342
DB 366 CTCACGAGCTGCTGTCGAAGGATTCACCTGGTTCAATGCTTTTATCAACGTAATCGAAT 425
QY 343 ACTCATGCTATCATCTCTATCTCGATTAACTACCTCTTTCGCGGATTAACGCGAGTTT 402
DB 426 AGCAATCATATCATCTCTATCTCGATTAACTACCTCTTTCGCGGATTAACGCGAGTTT 485
QY 403 CTTCTTGTGCTTCTAAGCTGAAGCGCGCAACTTTCACCTGATTATCATCTCTTTCGCT 462
DB 486 TTGCTTGTGCTGCTGTTAAGTTGAGACACAACTTGTACGGATTACATTATCTTTACGT 545
QY 463 TCAGACGATATCTCAAGGAAGAACGCGGTATCTTTGGGAGAAATGAGATCGAACTTCCTT 522
DB 546 TCTGATGATATGTCGAGAGAACCGCAAGCTTATGTTGGCAAGTGAATCGAACTTCCTA 605
QY 523 GGNACAAATTCAGCTCTTGTGATGTTAGTCAG-----ACCGGACGCGAGAGTNG 573
DB 606 GGAACGAAATTCATCTGTCTTGTGATGAAATCTGCTGCTTTCACCGGAGCCGGAAGTTG 665
QY 574 CAGAAGAGCGCTCTTCTTAATTTTCAACAAAGTTTTCACCTAGAGTTCTCTCAGGAAAGTTAC 633
DB 666 AGAAGAGCGGATCTTATATATCCGCAAGAGTTTTCAGCAAAAGTTCTCTTGGAGTTAT 725

QY 634 CCCATCGCTCACATTTTCATACGAGTTTAAACGCTTTAGGCTCTCGGGGACCGAGAAAGT 693
DB 726 CTTGCTGCTCATATACATATAGCTGATGATGCTTTAGGATCCGGGGACCAAGAAAGT 785
QY 694 CGTTGCTATCATGATACAAATACCTATGAGCATCGTGGAGTCGGGAGGAGTAGTAGCTTCA 753
DB 786 CAATGCTTTATGGACACAATACCTACAGCAACAATGGAGCCTCAAGGAGTAGTCTTCA 845
QY 754 ACATCATTAAGCTCTTTTTCAGTCGGTCATCACCAAGCTTTTAGTCTCTCACTCAAAACCA 813
DB 846 CCATCAGAGTTTCCCTTACTCGGTACTCGGTCAACCTTTATCCAGGTCTCAGTCAAAACCA 905
QY 814 TTGGCAGCTAATAGTCATCATGAGCACTCAGGCAACAACCTGGGAGATCCACCATTTG 873
DB 906 TTACGAGTA-----GCTCAAGCCACCTGAAGAAACACCATTA 944
QY 874 GTGCTGAGCAACAAGCTCCACGGTGGCATGAGCAAGTTACGTTGCTGTGCTTAAATTTTC 933
DB 945 GTGCTGAGCAACAAGACACACCGGTGGCCACGAGCAGCTACGCTGCTGCTTGAATTTTC 1004
QY 934 CATGCTCAGTCACTAGTCGCTTCCGTTAAGAACTTTTCAAGCAAGTTCCCTGTGAA 984
DB 1005 CATGCCCTGTTCACAGTAGCTGCTGAGGAACCTTTTCAGCTCGTGGCAGGAGCTAGC 1064
QY 985 AGTGACTGTGAAGCAGGCGCAGACATCTCAGA-----GGATCATACTCCAGTTT 1032
DB 1065 TGTGTCAGTGGCAGCGGAATGTACCGGAGAGGAGCGGAGCGGATTAATTTGAGTTT 1124
QY 1033 GGGAAAGTTGGGAAGCAGCATGTTTACCATGGATTATGATATCCGATTTCTGCGTTTCAA 1092
DB 1125 GGGAAAGTCGGGAAGATATGTTTACGATGATTATGATACCCGATCTCAGCTTTCCAG 1184
QY 1093 GCGTTTGTATCTGCTCAGCAGATTTTGAACCAAGTTCCCTGTGAA 1140
DB 1185 GCTTTTGCATTTGCTTGAGCAGCTTTGAGACTAGACTGCTTGTGAA 1232

RESULT 4
AAC37041
ID AAC37041 standard; DNA; 1295 BP.
XX AAC37041;
AC AAC37041;
XX 17-OCT-2000 (first entry)
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 15967.
XX Arabidopsis thaliana. genetic mapping; gene expression control;
XX Hybridization assay; protein identification; signal transduction pathway; metabolic pathway;
XX promoter; termination sequence; ss.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
XX 19-APR-1999; 99US-0130077P.
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.

PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134321P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.

PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 22-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.

```
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 29-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 50.1%; Score 571.2; DB 3; Length 1295;
Best Local Similarity 71.6%; Pred. No. 5.1e-177;
Matches 851; Conservative 0; Mismatches 268; Indels 69; Gaps 5;

QY 1 ATGACGTTCCGAAGTTTACTCCAGGAATGCGGTCTAGGCCACACCGTGTAGTTTACGCGC 60
DB |||||
DB 46 ATGACCTTACGTAGCTTAATCTCTTGAGATGCGTTCGAGACCGCATCGTGTGGTCCACGAC 105
QY |||||
DB 61 -----CGCGCTCAACCGCTAATAGTTTCAGACCTTTCAGCTGGTGG 102
DB |||||
DB 106 CTTGCGCGCGCGACGCTGCGGATTCACCTCTGTGTCATCGCAAGATTATCGTGTGUA 165
QY |||||
DB 103 GAGCTCCCGGAGGAGCTGCTTAGAGAAATCCGTATTAGGTTTGAGACTGTTGACCGCGGC 162
DB |||||
DB 166 GAGATTCTGAGAGCTTCTTAGGAGATTCGATTCTGTTGAGCGCGGACGCGTGGC 225
QY |||||
DB 163 GATTGCGCTGCGCGGAAACGFTGGTGTGTCGCGCGTGTGTCGTAAGTGTGGAGGATT 222
DB |||||
DB 226 GGATGCGCTGACGACGCGGTGGTGTGTCGCGCGTGTGTCGTCGCTGGCGGCTA 285
QY |||||
DB 223 CTCACCAAGAGATTGAGCTGCTGTAATCTCTCTAATTTGACTTTCCCTATCTCC 282
DB |||||
DB 286 CTTATGAACGAACCGCTGCTGTCCTGAGATCTCTTCTAAGTTGACTTCCCATCTCT 345
QY |||||
DB 283 CTCACGAGCTGCTGCTCAAGAGATTCTCTAGTTTCAATGCTTTTATAAAAGTAAATCGAAAT 342
DB |||||
DB 346 CTCACGACGCTGTCCAGGGGATTCACCTGTTCAATGCTTTTATCAAAAGTAAATCGAAT 405
QY |||||
DB 343 ACTCAATCGTATCATCTCTATCTCGGATTAACATCACTCTTTGACGGATTAACGGAGTTT 402
DB |||||
DB 406 ACGCAATCATATCATCTCTATCTCGGATTAACCAACTCTTTAACGGATGATGGGAAGTTT 465
QY |||||
DB 403 CTTCTTGTCTTCTTAAGCTGAGCGCGCAACTTTCGACTGATTACATCATCTCTTTGCGT 462
DB |||||
DB 466 TTGCTTGTCTGCTGTTAAGTTGAAGCACACAACTTGTACGGATTACATTTCTCTTTACGT 525
QY |||||
DB 463 TCAGACGATATCTCAAGAGAAACCAACGCTATCTTTGGGAGATGAGATCGAACTTCCTT 522
DB |||||
DB 526 TCTGATGATATGTCGAGAGAACCAAGCTTATGTTGGCAAGTGAGATCGAACTTCCTTA 585
QY |||||
DB 523 GGAAACAAATTCACCGTCTTTGATGTTAGTGTGTCAG-----ACCGAGACGCGAAGATG 573
DB |||||
DB 586 GGAAACGAATTCACCTGTCTTGTAGTAAATCTGCTCTTCAACGGGAGCGCAAGATTG 645
QY |||||
DB 574 CAGAAAGCGCTCTTCTAATTTTATCAAAAGTTTTCACCTAGACTTCTCAGGGAAGTTAC 633
DB |||||
DB 646 AGAAAGAGCGGATCTTATATAATCCGCAAAAGTTTTCAGCAAAAGTTTCTCTTGGAGTTAT 705
QY |||||
DB 634 CCCATCGCTCACATTTTCATACGAGTTTAAACGCTCTTAGGCTCTCGGGGACCGGAGAAATG 693
DB |||||
DB 706 CCTGTCGCTCATATCATATGAGCTGAATGTCTTAGGATCCCGGGGACCAAGAAAGATG 765
```

```
QY 694 CGTTGCATCATGGATACAAATACCTATGAGCATCGTGGAGTCGCGAGGAGTAGTAGTTCA 753
DB |||||
DB 766 CAATGCTTATGGACACATATACCTACAAGCACATGGAGCTCAAGGAGTAGTTCA 825
QY 754 ACATCCATAAAGCTCTTTTTCAGTGGTTCATCACAGTCTTTAGGTCTCACTCAAAACCA 813
DB |||||
DB 826 CCATCAGAGTTTCCCTTACTCGGTACTCGGTCAACCTTATCCAGGTCTCAGTCAAAACCA 885
QY 814 TTGGCAGTAATAGTGCATCATGTAGCGACTCAGGCAACAACCTGGGAGATCCACATTTG 873
DB |||||
DB 886 TTACGCAGTA-----GCTCAAGCCACCTGAAAGAAACACCATTA 924
QY 874 GTGCTGAGCAACAAGCTCCACGGTGGCATGAGCAGTTTACGTTGCTGTGCTTAAATTTT 933
DB |||||
DB 925 GTGCTGAGCAACAAGCAGACACCGGTGGCAGCAGCTACGCTGCTGTGCTTGAATTTT 984
QY 934 CATGTCGAGTCACAGTGGCTTCGGTTAAGAACTTTTACGCTTGTGGCAG-----TT 984
DB |||||
DB 985 CATGSCCGTGTACAGTAGCGTCACTGAAAGAACTTTTACGCTGCTGGCAGCAGGAGTAGC 1044
QY 985 AGTGACTGTGAAGCAGGCGCAGACATCTGAGA-----GGATCATACTCCAGTTT 1032
DB |||||
DB 1045 TGTGCGAGTGGCAGCGGAATGTCAACCGAGAGGCGAGCGGATTTATATTGCAAGTTT 1104
QY 1033 GGGAAAGTTGGGAAGGACATGTTTACCATGGATTATGGATATCCGATTTCTCGGTTTCAA 1092
DB |||||
DB 1105 GGGAAAGTCGGGAAGATATGTTTACGATGGATATGATGATACCCGATCTCAGCTTTCAG 1164
QY 1093 GCGTTTGTCTATCTGCTGAGCAGGTTTGAAGAACAGAAATTCGCTGTGAA 1140
DB |||||
DB 1165 GCTTTTGCATTTGCTTGGCAGGCTTTGAGACTAGATCGCTTTGTGAA 1212

RESULT 5
ADW64876
ID ADW64876 standard; cDNA; 1140 BP.
XX AC ADW64876;
XX DT 07-APR-2005 (first entry)
XX DE Arabidopsis thaliana TUBBY-like protein (TLP) 11 encoding cDNA.
XX KW Transgenic plant; salt tolerance; crop improvement; cold tolerance;
XX KW oxidative stress; drought resistance; TUBBY-like protein; plant;
XX OS Arabidopsis thaliana.
XX PH Key Location/Qualifiers
FT CDS 1..1140
FT /*tag= a
FT /product= "Arabidopsis thaliana TLP11 protein"
FT /partial
FT /note= "No stop codon"
XX US2005014266-A1.
XX 20-JAN-2005.
XX 21-JAN-2004; 2004US-00763042.
XX 21-JAN-2003; 2003US-0441380P.
XX (SINI-) ACAD SINICA.
XX Shaw J, Lai C;
XX WPI; 2005-080955/09.
DR P-PSDB; ADW64865.
DR GENBANK; AY046922.
XX
```


PI Hurban P;
XX WPI; 2003-110411/10.
XX Novel Arabidopsis thaliana nucleic acid useful for identifying homologous
PT or related genes, and to create genetically modified and transgenic
PT organisms, such as plant cells and plants.
XX
PS Claim 1; SEQ ID NO 879; 43pp; English.
XX
XX The invention describes an Arabidopsis thaliana nucleic acid (I). The
CC polypeptide (II) encoded by (I), transgenic plant (III) or genetically
CC modified cell (IV) are useful for screening a candidate agent for its
CC biological effect, by combining the candidate agent with (II), (III) or
CC (IV), and determining the effect of the candidate agent on (II), (III) or
CC (IV). (I) is useful for identifying homologous or related genes, for
CC producing compositions that modulate the expression or function of its
CC encoded protein, for mapping functional regions of the protein, in
CC diagnosis, for studying associated physiological pathways, for genetic
CC manipulation of cells, preferably plant cells, in screening assays of
CC various plant strains to determine the strains that are capable of
CC withstanding a particular disease or environmental stress, for enhancing
CC or inhibiting production of biosynthetic product in a plant, for
CC producing polypeptides, as probes for the detection of mRNA in biological
CC samples, to generate additional copies of (I), to generate ribozymes or
CC oligonucleotides, as single stranded DNA probes or as triple-strand
CC forming oligonucleotides, and to create genetically modified and
CC transgenic organisms, such as plant cells and plants. (II) or (III) is
CC useful for introducing or improving disease resistance and stress
CC tolerance in plants, screening biological active agents, e.g.,
CC fungicides, insecticides, etc., and for elucidating biochemical pathways.
CC (III) is useful as crops for their enhanced diseased resistance, enhanced
CC traits of interest, for screening programs, as crops which exhibit
CC enhanced tolerance to environmental stress, or to produce a factor. This
CC sequence represents a nucleic acid that may correspond to naturally
CC occurring Arabidopsis thaliana expressed sequences. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from the US patent office at
CC seqdata.uspto.gov/sequence.html?docID=999909770423
XX
SQ Sequence 396 BP; 97 A; 91 C; 89 G; 119 T; 0 U; 0 Other;

Query Match 34.7%; Score 396; DB 8; Length 396;
Best Local Similarity 100.0%; Pred. No. 1.7e-119; Indels 0; Gaps 0;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 160 GCGATTGGCGGCGGAAACGTGGTGGCTTGTGCGCGGCTTTGTGCTAGCTGGAGG 219
DB 1 GCGATTGGCGGCTGCGGCGAAACGTGGTGGCTTGTGCGCGGCTTTGTGCTAGCTGGAGG 60

QY 220 ATTCTCACAAGAGATTGAGCTGTTCTGTAATTCCTCTAAATGACTTTCCCTATC 279
DB 61 ATTCTCACAAGAGATTGAGCTGTTCTGTAATTCCTCTAAATGACTTTCCCTATC 120

QY 280 TCCTCAACGAGTCTGTCACAGAGATTCTTAGTTCAATGCTTTTATAAAGCTAATCGA 339
DB 121 TCCTCAACGAGTCTGTCACAGAGATTCTTAGTTCAATGCTTTTATAAAGCTAATCGA 180

QY 340 AATACTCAATCGTATCATCTATCTATCTCGGATTAACCTCTTTGACGGATGAGTCGAACTC 399
DB 181 AATACTCAATCGTATCATCTATCTATCTCGGATTAACCTCTTTGACGGATGAGTCGAACTC 240

QY 400 TTTCTTTCTGCTGCTTCTAAGCTGAAGCGCGCAACTTGACCTGATTAATCATCTCTTTTG 459
DB 241 TTTCTTTCTGCTGCTTCTAAGCTGAAGCGCGCAACTTGACCTGATTAATCATCTCTTTTG 300

QY 460 CGTTACAGAGATATCTCAAGAGAACCAACGCTATCTTTGGGAGATGAGTCGAACTTC 519
DB 301 CGTTACAGAGATATCTCAAGAGAACCAACGCTATCTTTGGGAGATGAGTCGAACTTC 360

QY 520 CTTTGAACAAAATTACCGGTCTTTTGATGGTAGTCAG 555
DB 361 CTTTGAACAAAATTACCGGTCTTTTGATGGTAGTCAG 396

RESULT 7
ADW16693
ID ADW16693 standard; cDNA; 1681 BP.
XX
AC ADW16693;
XX
DT 24-MAR-2005 (first entry)
XX
DE Eucalyptus grandis transcription factor cDNA TUBBY family Seq 421.
XX
KW gene; ss; plant; transcription; gene regulation; gene expression;
KW transgenic plant; drought resistance; disease resistance; salt tolerance;
KW cold tolerance; freezing tolerance; flowering; flavor enhancer;
KW flower color.
XX
OS Eucalyptus grandis.
XX
PN WO2005001050-A2.
XX
PD 06-JAN-2005.
XX
PF 07-JUN-2004; 2004WO-US017965.
XX
PR 06-JUN-2003; 2003US-0476189P.
XX
PA (ARBO-) ARBorgen LLC.
XX
PI Bloksberg LN, Bryant C, Connett MB, Emerson SJ, Frost MJ;
PI Forster RLS, Grigor M, Higgins C, Lasham A, Lund ST, Magusin A;
PI Phillips J, Puthigae S, Veerakone S, Westwood C, Gause K, Wood M;
XX
DR WPI; 2005-075542/08.
DR P-PSDB; ADW17479.
XX
XX New polynucleotides isolated from plants encoding transcription factors,
PT and polypeptides encoded by such polynucleotides, useful for regulating
PT gene transcription and gene expression.
XX
PS Claim 3; SEQ ID NO 421; 1265pp; English.
XX
XX This invention relates to novel isolated plant nucleic acid molecules, or
CC variants thereof, that encode transcription factors. Specifically, it
CC refers to transcription factor proteins that are capable of binding to
CC DNA in order to regulate gene transcription and gene expression in a
CC plants, in particular Eucalyptus grandis and Pinus radiata. The present
CC invention describes DNA constructs containing DNA encoding a
CC transcription factor that regulates the promoter, which is operably
CC linked to the desired nucleic acid to be expressed. It further provides
CC transgenic plants expressing a transcription factor that confers a trait
CC to the plant such as increased drought, salt or disease tolerance, height
CC change, enhanced cold/frost tolerance, enhanced color, health and
CC nutritional characteristics, as well as improved taste, starch
CC composition, flower longevity and germination, amongst others.
CC Accordingly, such plants that are successfully transfected with a DNA
CC construct can be characterized by a difference in flower color, petal or
CC leaf shape and size, aroma or plant height. This polynucleotide is a
CC plant transcription factor cDNA sequence of the invention.
XX
SQ Sequence 1681 BP; 381 A; 418 C; 464 G; 418 T; 0 U; 0 Other;

Query Match 34.0%; Score 387.4; DB 14; Length 1681;
Best Local Similarity 62.9%; Pred. No. 2.7e-116;
Matches 690; Conservative 0; Mismatches 386; Indels 21; Gaps 5;

QY 61 GCGGCTCAACCGCTAATAGTTTCAGACCTTTTCAGCTGGTCCGAGCTCCCGAGGAGCTG 120
DB 285 GCGGCGCGGAGCGAGACGGCTTGACCGGAGCTCTCTGGGCCCAACATCGCGAGAGCTG 344

QY 121 CTTAGAGAAATCCTGATTAGGGTTGAGACTGTTTACGGCGGCGGATTTGGCGCTCGCGCGCA 180
DB 345 CTGCGGAGGTTCTGTTGAAGATCGAGGCGTCCGGA---GGGGTCTGTCGCGCGGAAG 401

```
QY 181 AACGTGGTGGCTTGTGCGCGGCTTTGTCTGTAGCTGGAGGATTTCTACCAAGGAGATGTA 240
DB 402 AGCGTGGTGGCTTGTGCGCGGCTTTGTCTGTAGCTGGAGGATTTCTACCAAGGAGATCGTG 461
QY 241 GCTGTTCTCTGAATTTCTCTTAATTTGACTTTCCCTATCTCCCTCAAGCAGTCTGGTCCA 300
DB 462 AAGTCCCTGAGTCTCCGGAATCTTGAAGTTCCTCCATCTCGGTTAAACAGCCTGGGCT 521
QY 301 AGAGATTTCTAGTCAATGCTTTATAAAGTAATGAAATGAAATGAAATGAAATGAAATG 360
DB 522 AGGAATCTCTGTCAGTCTTTATCAAAAGAAACAGGTCCCAATCAGACGATCATCTC 581
QY 361 TATCTCGGATTAATCACTCTTTGACGGAATGAAAGTCTTTCTTCTGCTGCTTCTAAG 420
DB 582 TATGTCGGTTGACTAATGACTTACGATGACGGAAGTTCTTCTCGCTGCACGCAAG 641
QY 421 CTGAAGCGGCACTTGTCACTGATTAATCATCTCTTTGGTTTCAGACGATATCTCAAG 480
DB 642 TGTAGAGCCGCCACCTGCAAGATTAATGATCTCTCTGATGCTGATGATGCTCTAAG 701
QY 481 AGAAGCAACGCGTATCTTGGGAGATGAGATCGAACTTCTTGGAAACAAATTTACGGTC 540
DB 702 GGTAGCTGCACCTATGTTGMAAACTAAAGTCTAAATTTTCTTGGAAACCAAGTTCAATC 761
QY 541 TTTGATGGTAGTCAGAC----CGGAGCAGGAGATGAGAGAGCGCGCTCTTCTTAATTC 597
DB 762 TCTGATGGGCAACCACTCATCTCAGGAGCAAGATGCCAAAAGTCTCAACCAAGCTA 821
QY 598 ATCA-----AGTTTCACTAGAGTTCTCTCAGGGAAGTTTACCCCATCGCTCAATT 648
DB 822 GTCAGTTCAAAACAGGTTTCACTAGGCTTCAGCGGAACTATCCGGTGGCCCACT 881
QY 649 TCATACGAGTTAAACGCTTTAGGCTCTCGGGACCGAGAGAAATGCGTTGCTCATGAT 708
DB 882 TCTTATGAACTGATGTTGTTAGTTCAAGAGGACCAAGGAGATGCACTGTACTATGAT 941
QY 709 ACAATACCTATGAGATCTGAGTGGAGTGGAGGAGTAGTAGTTCAACATCCATAAGCTCT 768
DB 942 GCCATCCCTCGCTCAGCAATTTGAAGCTGGGGAGTTGCCCAACGAGACAGAGTTGCCT 1001
QY 769 TTTTCCAGTC---GGTCATCACAGTCTTTAGTCTCACTCAAAACCAATGGCAGTAAT 825
DB 1002 CTTAGCAGTTCGAGTCTTTCCCTCACTTCTTTTTCGAAACAAATCAAAATCGCGG 1061
QY 826 AGTGCATCATGATGAGCTCAGGCAACCACTGGGAGATCCCACTTGTGCTGAGCAAC 885
DB 1062 GAGAGTGTAGTTACGACCTCTGTTAGTCAGAAAGATGGAGCCTTGTGTTGAGAAAT 1121
QY 886 AAGCTCCACGGTGGATGAGCAGTTACGTTGCTGCTTAAATTTCCATGTTGAGTC 945
DB 1122 AAGGCTCTAGTGGGCAACCACTCAGTGTGTTGCTTGAATTTTTCATGGACGAGTA 1181
QY 946 ACAGTGGCTCGTTAAGAACTTACGTTGAGTGGCAGTAGTAGTCACTGTGAA---GCAGG 1002
DB 1182 ACAGTTGCTCGTCAAGAACTTTCAGTGTGTTGCTTCTCCGAAATGAACCTGCTGG 1241
QY 1003 CAGACATCTGAGAGGATCACTACCTCAGTTTGGAAAGTTGGGAGACATGTTTACATG 1062
DB 1242 CCCCAACATGAGAAATCATCTCCAGTTTGGAGGTTGGAGGATTTATTTACATG 1301
QY 1063 GATTTAGGATATCCGATTTCTGCGTTTCAAGCGTTTGTCTATCTGCTCAGCAGTTTGA 1122
DB 1302 GATTATCGGTATCCCATCTCAGCAATTCAGGCTTTTGCAATCTGCTCAGCAGCTTTGAC 1361
QY 1123 ACAGAAATTCCTGTGA 1139
DB 1362 ACAAGATTCGCTGTA 1378
```

RESULT 8

ADX13792

ID ADX13792 standard; cDNA; 1629 BP.

```
XX AC ADX13792;
XX DT 21-APR-2005 (first entry)
XX DE Plant full length insert polynucleotide seqid 8367.
XX KW plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
KW protein content; gene; ss.
XX OS Unidentified.
XX PN US2004034888-A1.
XX PD 19-FEB-2004.
XX PF 28-APR-2003; 2003US-00425114.
XX PR 06-MAY-1999; 99US-00304517.
XX PR 05-NOV-2001; 2001US-00985678.
XX PA (LIU/) LIU J.
XX PA (ZHOU/) ZHOU Y.
XX PA (KOVA/) KOVALIC D K.
XX PA (SCRE/) SCREE S E.
XX PA (TABR/) TABASKA J E.
XX PA (CAOY/) CAO Y.
XX PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX WIPI; 2004-180133/17.
XX PT New recombinant DNA construct, useful for improving plant tolerance to
XX PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
XX PT pests, for conferring increased resistance to plant disease, or for
XX PT improving yield.
XX PS Claim 1; SEQ ID NO 8367; 15pp; English.
XX CC The invention describes a recombinant DNA construct comprising a
XX CC polynucleotide consisting of a sequence encoding an amino acid sequence
XX CC available in electronic form from the US patent office at
XX CC ftp.sequata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
XX CC of the invention are also useful in physical arrays of molecules and as
XX CC plant breeding markers. The recombinant DNA construct is useful for
XX CC improving plant tolerance to cold, heat, drought, herbicides, extreme
XX CC osmotic conditions, pathogens or pests, for manipulating growth rate in
XX CC plant cells by modification of the cell cycle pathway, for conferring
XX CC increased resistance to plant disease, for producing galactomannan,
XX CC lignin or plant growth regulators, for increasing the rate of homologous
XX CC recombination in plants, for improving yield by modification of
XX CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
XX CC or by providing improved plant growth and development under at least one
XX CC stress condition or for modifying seed oil or protein yield and/or
XX CC content. This sequence represents a plant full length insert
XX CC polynucleotide that can be used in the recombinant DNA construct of the
XX CC invention.
XX SQ Sequence 1629 BP; 459 A; 369 C; 363 G; 438 T; 0 U; 0 Other;
```

Query Match 32.3%; Score 367.8; DB 13; Length 1629;
Best Local Similarity 63.2%; Pred. No. 7.9e-110;
Matches 674; Conservative 0; Mismatches 367; Indels 26; Gaps 6;

QY 95 GCTGTCGAGCTCCGGAGAGCTCTTAGAGAAATCTGATTAGGTTGAGACTGTTG 154
DB 156 GTTGGGCCAACATGCCCAAGAGCTTCTCCGAGAGTCTCTCCGAATCGAG---GCCT 212

Qy	155	ACGCGCGGAGATTGGCCGTGCGCGGCGAAACGTGGTGGCTTGTGCGCGGCTTTGTGCTAGCT	214
Db	213	CCGAGGACACGTGTGCGCGCGCGAAGAGCGTCTCTCTGCGCGCGGCTGTGCGCGAGCT	272
Qy	215	GGAGGATTTCTACCAAGCGAGATTCTAGCTGTTCCTCTGAATTCCTCTCTTAAATTTGACTTTCC	274
Db	273	GGAGACATATCACCAAGACATTTGTCAAACGCCCGAACTCTCTCTCCAAGATCACCTTCC	332
Qy	275	CTATCTCCCTCAAGCAGTCTGGTCCAAGAGATTTCTTAGTTCAATGCTTTATAAAGCTA	334
Db	333	CCATTTCTGTATAACAGCTGGCCCAAGGAGAACTCTCTACGGTGTCTTTATAAAGCGCA	392
Qy	335	ATCGAATATCTCAATCGTATCATCTCTATCTCGGATTAATCTACCTCTTTGACCGATAACG	394
Db	393	ACGGTCCACCCAAACATATAATTGTTCCTCAGTTTAAACAGTACCTTAGCTGAAGTG	452
Qy	395	GGAGTTTCTTTCTTGTGCTTCTTAAGCTGAAGCGCGAACTTGCACTGTATTATCATCATCT	454
Db	453	GGAGTTTCTTTCTTGTGCGCAACATGACAGCTCCAACTGACAGATTATATCATCT	512
Qy	455	CTTTGCGTTTCAGACGATATCTCAAAGAGAGCAACGGTATCTTTGGGAGAAATGAGATCGA	514
Db	513	CTCTTGACGCGAGATGATATGTCAAAGGAAAGCACTCTATGTGTGGANAATAAGATCTAA	572
Qy	515	ACTTCTTTGGAACAAAATTCACGGTCTTTGATGTAGTTCAGAC---CGGAGCAGCGAAGA	571
Db	573	ACTTCTTTGGAACAAAGTTCAATCTATGATACCGAGCTGCCTCATCAGGAGCAAGA	632
Qy	572	TGCAGAGAGCGGCTCTTC-----TAATTTTCATCAAAGTTTTCACCTAGAGTTCTCTC	622
Db	633	TTATGAAGAGTCGCTCCACAGGCTGGTGAATCTTAAAGCAAGTTTCAACCAAGGTTCTCTA	692
Qy	623	AGGCAAGTTACCCCATCGCTCACATTTTCATACAGAGTTAAAGCTCTTAGGCTCTCGGGAC	682
Db	693	CAGGCAACTATTCAGTGCCCTATTTTCATATGAATTTGATGTGCTTAGGCTCAAGGGGTC	752
Qy	683	CGAAGAAGATGCGTTTGATCATATGATACAATCTATAGCATCGTGGAGTCGCGAGGAG	742
Db	753	CTAGGAAATGCAATTGTGTATGATTTCCATTCTCTGCCACTGCAATTTGAACCTGGAGGTG	812
Qy	743	TAGTAGCTTCAACATCCATAGCTCTTTTCCAGTCGGTCATCACCAAGTCTTTTAGGTCCTC	802
Db	813	TAGCACCTACACAGAC--TGAGTTTCTCTTAAACAATAGACATGTTTCTCTCAITCCC	870
Qy	803	ACTCAAAACCAATTCGCGAGTAATAGTGCATCATGTAGCGACTCAGG-----CAACAACC	856
Db	871	CTTTTCTTTCGATCAAAATCAAACTGTGCGGAAATTCATCTCTGGACCATTTGGTTGATC	930
Qy	857	TGGAGATCCACCAATTGGTGTGAGCAACAAGCTCCACGGTGGCATGACGAGTTAGCTT	916
Db	931	AAAAGSATTGGGATGCTAGTGTGTGAANAACAAGGGCCCCCTAGGTGGCATGAGCAGCTGCAAT	990
Qy	917	GCTGGTGTCTAAATTTTCATGGTCCAGTCAAGTGGGCTTCGGTTTAAAGAACTTTTCAGCTTG	976
Db	991	GTTCGTGCTGAACTTTTCATGGAACGGGTGACAAATGGCTCAGTTTAAAACTTTTCAGCTGG	1050
Qy	977	TGGCAGTTTAGTG---ACTGTGAACGAGGGCAGACATCTGAGAGGATCATACTCAGTTTG	1033
Db	1051	TGGCTTCTGCAGAAAAACGGACCTCTGTGACCCAGAACACGATAAGATCATCTCTCCAAATTG	1110
Qy	1034	GGAAAGTTGGGAAGGACATGTTTACCAAGTAATGGATATCCGATTTCTGCGTTTCCAG	1093
Db	1111	GAAAGTTGGGAAGGATTTGTTTACATGGATTTACCGGTACCTTATCTCGGCATTTTCAGG	1170
Qy	1094	CGTTTGTCTATCGCTCAGCAGTTTGTAAACACAGAAATTCGCTGTGAA	1140
Db	1171	CATTGTGAATCTGCTCAGCAGTTTGTGATCAAGAAATGCTGTGTGAA	1217

XX	ADW16694;
AC	
XX	
XX	24-WAR-2005 (first entry)
XX	
DE	Eucalyptus grandis transcription factor cDNA TUBBY family Seq 422.
XX	
XX	gene; ss; plant; transcription; gene regulation; gene expression;
KW	transgenic plant; drought resistance; disease resistance; salt tolerance;
KW	cold tolerance; freezing tolerance; flowering; flavor enhancer;
KW	flower color.
XX	
OS	Eucalyptus grandis.
XX	
WO	WO2005001050-A2.
XX	
XX	06-JAN-2005.
PD	
XX	
PF	07-JUN-2004; 2004WO-US017965.
XX	
XX	06-JUN-2003; 2003US-0476189P.
PR	
XX	
PA	(ARBO-) ARBORGEN LLC.
XX	
XX	Bloksberg LN, Bryant C, Connett MB, Emerson SJ, Frost MJ;
PI	Forster RLS, Grigor M, Higgins C, Lasham A, Lund ST, Magusin A;
PI	Phillips J, Puthigae S, Veerakone S, Westwood C, Gause K, Wood M;
PI	
WPI	WPI; 2005-075542/08.
DR	
XX	P-PSDB; ADW17480.
XX	
PT	New polynucleotides isolated from plants encoding transcription factors,
PT	and polypeptides encoded by such polynucleotides, useful for regulating
PT	gene transcription and gene expression.
XX	
PS	Claim 3; SEQ ID NO 422; 1265pp; English.
XX	
CC	This invention relates to novel isolated plant nucleic acid molecules, or
CC	variants thereof, that encode transcription factors. Specifically, it
CC	refers to transcription factor proteins that are capable of binding to
CC	DNA in order to regulate gene transcription and gene expression in a
CC	plants, in particular Eucalyptus grandis and Pinus radiata. The present
CC	invention describes DNA constructs containing DNA encoding a
CC	transcription factor that regulates the promoter, which is operably
CC	linked to the desired nucleic acid to be expressed. It further provides
CC	transgenic plants expressing a transcription factor that confers a trait
CC	to the plant such as increased drought, salt or disease tolerance, height
CC	change, enhanced cold/ frost tolerance, enhanced color, health and
CC	nutritional characteristics, as well as improved taste, starch
CC	composition, flower longevity and germination, amongst others.
CC	Accordingly, such plants that are successfully transfected with a DNA
CC	construct can be characterized by a difference in flower color, petal or
CC	leaf shape and size, aroma or plant height. This polynucleotide is a
CC	plant transcription factor cDNA sequence of the invention.
XX	
SO	Sequence 2586 BP: 617 A; 577 C; 690 G; 702 T; 0 U; 0 Other

209	GTAGCTGAGGATTTCTCACCAAGAGAGATTGTAGCTGTTCTCTGGAATTTCTCTCTAAATTTGA	268
QY		
478	GGAACTGGAGGGAGATCATGAAGGAAATTTGTGAAACACACGGAGGGTGTCTGGCAAGTTGA	537
DB		
269	CTTTTCCCTATCTCCCTCAAGCAGTCTCGTCCCAAGAGATTCTCTAGTTTCAAATGCTTTTATAA	328
QY		
538	CAITTCANATTTTGTTGAGCAGCGCTGGTCCAGGGACTCCCTTCTTCAGTGTATTATTTA	597
DB		
329	AACGTAAATCGAAATPACTCAATCGPATCATCTCTATCTCGGATTAACTACCTTTTGTGACGG	388
QY		
598	AACGCAATCGCAGCAACCAACATATTATCTCTACCTCAGCTTAGATCAAGCTTTTAAATG	657
DB		
389	ATAACGGGAGTTTCTTCTTCTGCTCTTAAGCTGAAGCGGCAACTTTCGACTGCAATTACA	448
QY		
658	AAGATGGCAAAATTTCTTCTCGCTGCACGTAAAGCTCCGGCGCCCTACTTGCACAGACTACG	717
DB		
449	TCATCTTTTGGCTTTCAGACGATATCTCAAAGAGAACAACGCGTATCTTTGGGAGAGATGA	508
QY		
718	TCATCTTTTAAATCCGATGACATGTCCAAAGGAGTAGCACATATTATTTGGAANAATTAA	777
DB		
509	GATCGAACTTCTTTGGAAACAAAATTCACGGTCTTTTGATG-----GTAGTCAGA	556
QY		
778	GGTCAAACTTTCTGGGCACGAAGTTTCACAGCTTTTGATGCACAGCTCCCAAGTCTCGCT	837
DB		
557	CCGAGCAGCGAAGATGCAGAGACGCGCTCTTCTAAATTTTCATCAAAAGTTTCACCTAGAG	616
QY		
838	CCAGAGTCATGAAAAGCCGGTCGACACAGGCTAGTGAACCTTTTAAAGCAAGTCTCCCGGAA	897
DB		
617	TTCTCTCAGGGAAGTTTACCCCATCGCTCACATTTTCATACGAGTTTAAACGTCTTAGGCTCTC	676
QY		
898	TTCCCGCTGGAAATTTACCTTCAGTGCACATATCTTATGAGTTGAAATGTCTTTGGGATCAA	957
DB		
677	GGGGACCGAGAAGAAATCGTTTGCATCATTGGATACAATACCTTATGAGCATCGTGGAGTCGC	736
QY		
958	GGGGACCGAGAAGAAATGCAATGTGTCTTGGAGGCCATACCTGCGCTTCTCAATTCGAGCCAG	1017
DB		
737	GAGGAGTAGTAGTTTCAAATCTCAATAGCTCTTTTTCAGTTCGGTCAATCACCAGTCTTTTA	796
QY		
1018	GAGGAGTTTGCCCGACACAGACTGAAATTTGTGTCCAGCACTCTCGACACATTTTCCCTCAT	1077
DB		
797	GGTCTCACTCAAACCAATTCGGCAGTAATAGTGCATCATGTAGCGACTCAGGCAACAACCC	856
QY		
1078	TCCTCTTTTTCAGATCTAAATTCGACGGCGGTAGAGACTTTGAGTCTGGGAGACCAATGTC	1137
DB		
857	T-----GGGAGATCCACCAATTTGGTGTCTGAGCAACAAAGCTCCACGGTGGCATGAGCAGTT	911
QY		
1138	TGCTCCGAAGATCGGATGCTGTATTGAAGAACAAAGTCTCCGAGGTGGCAAGCAACT	1197
DB		
912	ACGTTGCTGTGCTTAAATTTCCATGTGTCAGTCAAGTGGCTTCGGTTTGAAGAACTTTTCA	971
QY		
1198	CCAGTGTGCTGTCTAAATTTTTCATGACGGGTTTACTTGTGCTTTCAGTGAAGAACTTTCCA	1257
DB		
972	GCTTTGTGGCAGTTAGTGACTGTGAACAGGGCAGACA-----TCTGAGAGGATCATACTCCA	1028
QY		
1258	GTTTGGTGGCTTCTCCAGAACAGAGGTTCGCGGTACAGAGCATGAGAACGTTATCTCTCCA	1317
DB		
1029	GTTTGGGAAAGTTGGGAAGCACATGTTTACATGGATTTATGGATATCCGATTTCTCGGTT	1088
QY		
1318	GTTTGGGAAAGTTGGGAGGAGTATTATTCANATGGATATTCANATATCTATCTCAGCTTT	1377
DB		
1089	TCAAGCGTTTGCCTATCTGCTGACAGCACTTTTGAACCCAGAAATTCGCTGTGAA	1140
QY		
1378	CCAATCATTTGGCATAATGTCTCGACAGTTTTCGACACTTCGAATCGCTGTGAA	1429
DB		

RESULT 10

RESULT I
ADW64868

ADW64868
ID ADW64868 standard; cDNA: 1218 BP.

AC ADW64868:

XX
XX
'00050WNY

- 2 -

DT 07-APR-2005 (first entry)

XX
XX

DE	Arabidopsis thaliana TUBBY-like protein (TLP) 3 encoding cDNA.
XX	
KW	Transgenic plant; salt tolerance; crop improvement; cold tolerance;
KW	oxidative stress; drought resistance; TUBBY-like protein; plant;
KW	chromosome 2; gene; ss.
XX	
OS	Arabidopsis thaliana.
XX	
XX	Key Location/Qualifiers
FH	1. .1218
FT	/*tag= a
FT	/product= "Arabidopsis thaliana TLP3 protein"
FT	/partial
FT	/note= "No stop codon"
XX	
FN	US2005014266-A1.
XX	
XX	20-JAN-2005.
PD	
XX	
XX	21-JAN-2004; 2004US-00763042.
PF	
XX	
XX	21-JAN-2003; 2003US-0441380P.
PR	
XX	
PA	(SINI-) ACAD SINICA.
XX	
PI	Shaw J, Lai C;
XX	
XX	WPI; 2005-080955/09.
DR	P-PSDB; ADW64857.
DR	GENBANK; AY045774.
XX	
XX	New isolated Arabidopsis TUBBY-like proteins, useful for producing
PT	transgenic or transformed cells or plants having higher tolerance to
FT	salt, chilling, pathogens, oxidative stress, or water-deficit.
FT	
XX	
PS	Claim 7; SEQ ID NO 14; 43pp; English.
XX	
XX	The invention relates to Arabidopsis TUBBY-like proteins (TLP) and their
CC	corresponding nucleic acid sequences. The invention also relates to a
CC	transformed cell or a transgenic plant containing TLP nucleic acid and a
CC	transformed cell or a transgenic plant lacking TLP nucleic acid. The TLP
CC	polypeptides and nucleic acids are useful for producing transgenic or
CC	transformed cells or plants having higher tolerance to salt, chilling,
CC	pathogens, oxidative stress or water-deficit. The present sequence is the
CC	Arabidopsis thaliana TLP3 encoding cDNA. The Arabidopsis thaliana TLP3
CC	gene is located at chromosome II.
XX	
XX	Sequence 1218 BP; 320 A; 297 C; 283 G; 318 T; 0 U; 0 Other;

Query Match	30.1%;	Score 342.6;	DB 14;	Length 1218;
Best Local Similarity	59.8%;	Pred. No. 1.4e-101;		
Matches 677;	Conservative 0;	Mismatches 429;	Indels 27;	Gaps 5
Qy	32	GGTCTAGGCACACCGTGATGTTTCACGCGCGGCTCAACCGCTAATAGTTTCAGAGCCCTT	91	
Db	89	GATCCAGTCTCAACGTTGTTTCAGGATACTTCTGTTCTCTGTGATGCTTTTCAAGCAGA	148	
Qy	92	TCAGCTGTCGGAGCTCCCGGAGAGCTGCTTAGAGAAATCCTGATTAGGCTTGAGACTG	151	
Db	149	GTGTCCTGGCTAGTATGCTCCCGAGCTCTCCGAGAGATGTTCTTATGAGGATTGAGCAAT	208	
Qy	152	TTGACGCGCGCATTTGGCCGTCGCGCGGAAACGTGTGGTCTGTGCGCGGCTTTCCTGTA	211	
Db	209	CCGNAGACAC---TTGGCGGCTTAGGAAAATGTTGTTCTTGGCTGTGCTGCAGGA	265	
Qy	212	GCTGGAGATTCTCACAAGAGAGATTGAGCTGTTTCCTGAAATCTCTCTAAATGACTT	271	
Db	266	ACTGGCGAGAAATCGTCAAGAGATCGTCAGAGTTCCTGAGCTTCTAGCAAACTCACTT	325	
Qy	272	TCCCTATCTCCCTCAAGAGTCTGGTCCAAGAGATTCTCTAGTTCAAATGCTTTATAAAC	331	
Db	326	TTCTATCTCCCTCAAAACAGCGGGTCTAGAGATCACTTGTTCAAATGCTATATATGA	385	

QY 332 GTAATCGAAATACTCAATCGTATCATCTCTATCTCGATTAA---CTACCTCTTTGACGG 388
Db 386 GAAACCGGAGCAATCAAACCTACTACTATCTATATACCTCGGTTAAACCAAGCAGCTTCAAATG 445
QY 389 ATAACGGGAAGTTTCTTCTTGTCTGCTTTCTAAAGTGAAGCGCAACTTGTCACTGATTACA 448
Db 446 ATGATGGAAGTTTCTTCTTGTCTGCTTCTCAAGAGGTTTTCGGAGGCCAACTTGTCACTGACTACA 505
QY 449 TCATCTCTTTGGTTTCAGACGATATCTCAAAGAGCAACGGTATCTTTGGGAGATGA 508
Db 506 TCATCTCTTAAACTGCGATGATGTCTCTCGAGGAGCAATACCTATATATCGGAAGCTTTA 565
QY 509 GATCGAACTTCTTGAACAAATTCACGGTCTTTGATGTTAGTGTAGTCAGACCGGAGCAG--- 565
Db 566 GATCTTAACTTTCTGGGACCAAGTTCACTGTCTATGACGCTCAGCGAGCAATCTGGNA 625
QY 566 -----CGAAGATGCAAGAGCGCTCTTCTTAATTTTCATCAAGTTTCACTTAGAG 616
Db 626 CTCAGGTTACCAAGACCCGTTCAAGCAGACTTCTCAGTTTGAACCAAGTGAGCCCGAGAA 685
QY 617 TTCCTCAGGAGTTTACCCCATCGCTCACAATTTTCATACGAGTTAAACGTTTAGGCTCTC 676
Db 686 TTCCATCTGGCAATATCTCTGTAGCACAATATCTCATATGAGCTTAAACGTTTGGGTCCA 745
QY 677 GGGGACCGAGGAAGATGCGTTGCATCATGATACATACCTATGAGCATCGTGGAGTCGC 736
Db 746 GAGGACCGAGGAGGATGCGATGTGTCTATGATGCCATCCCTGCATCAGCTGTAGAACCTG 805
QY 737 GAGGAGTAGTAGTTTCAACATCCATAAGCTCTTTTTTCCAGTC-----GGTCATCACCGAG 790
Db 806 GAGGAACAGCTCCAACTCAGAGGAACTTGTCTCATAGCAATCTTGATAGTTTCCCGCTCAT 865
QY 791 TCTTTAGTCTCACTCAAAACCATTCGCGAGTAATAGTGCATCATGTAGGACTCAGGCA 850
Db 866 TCTCTCTTCTCAGTGGAAATCAATTCGTGCGAGAGTCTCCCTTCTGTCATCATCTG 925
QY 851 ACRACTGGGAGATCCACCATGCTGCTCGAGCAACAAAGCTCCACGGTGGCATGAGCACT 910
Db 926 CTGCTCAGAAGGAAGACTGCTTGTCTGAAACAAAGCCGACATGGCAGCAACAGC 985
QY 911 TACGTTGCTGTGCTTAAATTTTCCATGCTGCGAGTCACAGTGGCTTCGGTTTAAAGAACTTTC 970
Db 986 TCCAGTGTCTGTGCTCACTTCAATGGGAGAGTCACAGTGTCTTCCGTCAAAAACCTTTC 1045
QY 971 AGTTTGGCAGTTAGTACTGTGAA---GCGGGCAGACATCTGAGAGATCATACTCC 1027
Db 1046 AGCTGTGTGCTCTCCTGGAATGAGACCTGCGAGGACCTGAGCAGCAAAACGTTGATTTCC 1105
QY 1028 AGTTTGGGAAAGTTGGGAAGGACATGTTTACCATGATTATGATATCCGATTTCTGCGT 1087
Db 1106 AGTTTGGAAAGTCGGAAGAGATGTGTTCAAATGGATATCATAGTACCTTATCTGCGCT 1165
QY 1088 TTCAGGCTTTGCTATCTGCTCAGCAGATTTTGAACCCAGAAATGCTGTGAA 1140
Db 1166 TCCAGGCTTACCAATTTGCTCAGCAGTTTTCGACCAAGATGATGTGAA 1218

RESULT 11
ADN73404
ID ADN73404 standard; cDNA; 1221 BP.
XX
AC ADN73404;
DT
15-JUL-2004 (first entry)
XX
Thale cress cDNA upregulated in E2Fa/Dpa expressing plants SeqID 1299.
DE
XX
gene; ss; plant; transgenic; E2Fa/Dpa transcription factor;
KW growth regulator; animal feed product; thale cress;
KW cell wall biosynthesis; nitrogen metabolism; carbon metabolism.
XX
OS Arabidopsis thaliana.
XX

PN W02004035798-A2.
XX 29-APR-2004.
XX 20-OCT-2003; 2003WO-EP011658.
XX 18-OCT-2002; 2002EP-00079408.
XX (CROP-) CROPDESIGN NV.
PA Inze D, De Veylder L, Vlieghe K;
PI WPI; 2004-348466/32.
XX P-PSDB; ADN73405.
DR
XX
PT Altering plant characteristics, useful for producing plants for enzyme or
PT pharmaceutical production comprises modifying in a plant, expression of
PT one or more nucleic acids and/or modifying level or activity of one or
PT more proteins.
XX
PS Claim 1; SEQ ID NO 1299; 134pp; English.
XX
CC This invention relates to a novel method for altering one or more plant
CC characteristics. Specifically, it refers to identifying genes that are up
CC - or down-regulated in transgenic plants overexpressing the heterodimeric
CC E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to
CC alter plant characteristics accordingly. The present invention describes
CC generating transgenic plants for the production of growth regulators,
CC enzymes, therapeutics, pharmaceuticals and animal feed products, where
CC the altered plant characteristics are selected from increased yield or
CC biomass, enhanced survival capacity, stress tolerance, plant architecture
CC or physiology, altered endoreplication and/or altered photosynthesis,
CC each relative to the corresponding wild type plants. Accordingly, these
CC sequences can also be useful as positive or negative selectable markers
CC during transformation of cells or tissues. The identified genes play a
CC role in a variety of biological processes such as DNA replication, cell
CC wall biosynthesis, nitrogen and/or carbon metabolism or they function as
CC transcription factors. This polynucleotide sequence is thale cress cDNA
CC upregulated 1.3 fold or more in plants overexpressing the E2Fa/Dpa
CC transcription factor, given in an exemplification of the invention.
XX
SQ Sequence 1221 BP; 321 A; 297 C; 284 G; 319 T; 0 U; 0 Other;
Query Match 30.1%; Score 342.6; DB 12; Length 1221;
Best Local Similarity 59.8%; Pred. No. 1.4e-101;
Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;
QY 32 GGTCTAGGCCACACCGTGTAGTTTACGCGCGCCTCAACCGCTAATAGTTTCAGACCCCTT 91
Db 89 GATCCAGGCTCTCAACGTTGTTTTCAGGATACCTTCTGTTCTCTTTCAGACGAGA 148
QY 92 TCAGTGTGTCGAGCTCCGCGAGGAGCTGTTAGAGAAATCTCTGATTAGGTTGAGACTG 151
Db 149 GCTGCTGGGCTAGTATGCTCCGAGAGCTCTCAGAGATGTTCTTATGAGGATGAGCAAT 208
QY 152 TTGACGGGGGATTTGGCGTCGCGGAAACGTTGGTGGCTTGTGCGCGCTTTTTCGTA 211
Db 209 CCGAAGACAC---TTGGCGCTCTAGGAAATATGTTTCTTTCGCGCTGTTCTGAGGA 265
QY 212 GCTGGAGGATTTCTCACCAGGAGATTTGAGCTGTTCTCGAATTTCTCTCTAAATTCAC 271
Db 266 ACTGGCGAAGATTCGTCAGAGATCGTCAGAGTTCCTGAGCTTCTACGAACTCATT 325
QY 272 TCCCTATCTCCCTCAAGCAGTCTGGTCCAGAGATTTCTTAGTTCATGCTTTATATAAC 331
Db 326 TTCTCTATCTCCCTCAACAGCGGGTCTCAGAGATCACTTGTTCATGCTATATATGA 385
QY 332 GTAATCGAAATACTCAATCGTATCATCTCTATCTATCTCGATTAA---CTACCTCTTTGACGG 388
Db 386 GAAACCGGAGCAATCAAACCTACTACTATATCTCGGTTAAACCAAGCAGCTTCAAATG 445
QY 389 ATAACGGGAAGTTTCTTCTTGTCTGCTTCTTAAGCTGAAGCGCGCAACTTGCACCTGATTACA 448

Db 822 TAACTCTTGGGAACCAATTTACTATCTATGATGCCATCCACCTTATGCTGGAGATGT 881
QY 566 -----CGAAGATGCGAAGAGCGGCTCTCTTAATTTTCATCAAAAGTTTACCTAGAGTTCC 620
Db 882 GATTTTCAAAAGGCTCAGTCTGCGACGTGTGATTTGGTTCAAAACCATTTTATCCCCAGGATACC 941
QY 621 TCAGGGAAGTTACCCATCGCTCACATTTTCATACGAGTTTAAACGTTTAAAGCTCTCGGG 680
Db 942 TCGTGGGAATATCCAGTCTCAATATTTTCATATGAATTTGAACGTTTGGGTTCTAGGG 1001
QY 681 ACCGAGAAGATGCGTTGTCATCATGATGATACAATACCTATGAGCATCGTGGAGTCGCGAG 740
Db 1002 CCGAGAAGATGCACTGCGCTATGATTCATCCCTGTATCAGCAATTTGAACAGGAG 1061
QY 741 AGTAGTAGTTCAACATCATAGTCTTTTTTCCAGTCGGTATCACCAAGTCTTTAGGTC 800
Db 1062 AACAGTCCGACGCACTGAAATTTCTTTGAGTTACCATGAATCTTTTCAATCAATTC 1121
QY 801 TCACTCAAAACCATTCGCGAGTAATAGTGCATCATGTAGCGACTCAGGCAACACCTGG 860
Db 1122 CTTTTTCAATCAAAATCAGTTTGGAGCAATTAATTCATGCTTCAATTACTCACTCAAAA 1181
QY 861 AGATCCACCATTTGCTGTCAGCAACAAAGCTCCACGGTGGCATGAGCAGTACGTTGCTG 920
Db 1182 TGAAGCAAACTGGTGTGTGAGNACAAGTCCCGAGGTGGCAGACATCTGCGAGTCTG 1241
QY 921 GTGCTTAAATTTCCATGTCGAGTCACAGTGGCTTCGGTTAAGAACTTTACGCTTTGGC 980
Db 1242 GTGCTTAACTTCCATGGAGCGGTAAAGCTTGCATCGGTAAAGAACTTCCAGCTGGTGGC 1301
QY 981 AGTTAGTGACTGTGA---AGCAGGGCAGACATCTGAGAGGATCATACTCAGTTTGGGAA 1037
Db 1302 TTCAATGAGAGTAACCCAAACCAACGACGATGATGTATCTCCAGTTTGGTAA 1361
QY 1038 AGTTGGGAAGGACATGTTTACCATGATTTATGGATATCCGATTTCTGCGTTTCAAGCGTT 1097
Db 1362 AGTTGGGAAGGACATGTTTACCATGATTTATGGATATCCGATTTCTGCGTTTCAAGCGTT 1421
QY 1098 TGCTATCTGCTGAGCAGTTTGAACACCAAGAAATTCGCTGTGAA 1140
Db 1422 TGGATATGCTGAGCAGCTTTGACACCAAAATAGGTGGAA 1464

RESULT 13
ID ABL93451/C
XX ABL93451 standard; cDNA; 469 BP.
AC ABL93451;
XX
XX 10-JUN-2002 (first entry)
DT
XX Arabidopsis thaliana nucleic acid sequence Ref:2027216 SEQ ID NO:216.
DE Arabidopsis thaliana; insecticide; fungicide; plant; mapping; diagnosis;
KW Arabidopsis thaliana; insecticide; gene; ss.
XX Arabidopsis thaliana.
OS Arabidopsis thaliana.
XX
XX US2002023280-A1.
EN
XX 21-FEB-2002.
PD
XX 26-JAN-2001; 2001US-00770444.
PF
XX 27-JAN-2000; 2000US-0178502P.
PR
XX (GORL/) GORLACH J.
PA (ANY/) AN Y.
PA (HAMI/) HAMILTON C M.
PA (PRIC/) PRICE J L.
PA (RAIN/) RAINES T M.
PA (YUY/) YU Y.

PA (RAME/) RAMEAKA J G.
PA (PAGE/) PAGE A.
PA (MATH/) MATHW A V.
PA (LEDF/) LEDFORD B L.
PA (WOES/) WOESSNER J P.
PA (HAAS/) HAAS W D.
PA (GARC/) GARCIA C A.
PA (KRIC/) KRICKER M.
PA (SLAT/) SLATER T.
PA (DAVI/) DAVIS K R.
PA (ALLE/) ALLEN K.
PA (HOFF/) HOFFMAN N.
PA (HURB/) HURBAN P.
XX
PI Gorlach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y;
PI Rameaka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;
PI Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N;
PI Hurban P;
XX WPI; 2002-267486/31.
XX
XX New Arabidopsis thaliana nucleic acid, for identifying homologous genes,
PT producing compositions that modulate the expression or function of its
PT encoded protein, and mapping functional regions of a protein.
XX
PS Claim 1; SEQ ID NO 216; 44pp; English.
XX
XX The present invention describes an Arabidopsis thaliana nucleic acid (1)
CC comprising a sequence capable of hybridising under stringent conditions
CC to a sequence (S1) selected from any one of the 999 sequences given in
CC ABL93236 to ABL94234. (1) have insecticide and fungicide activities, and
CC they can be used as protein expression modulators. (1) can be used in
CC identifying homologous or related genes, in producing compositions that
CC modulate the expression or function of their encoded proteins, mapping
CC functional regions of the proteins, and in studying associated
CC physiological pathways. (1) can also be used: (1) for the genetic
CC manipulation of cells, particularly plant cells; (2) in screening assays
CC of various plant strains to determine the strains that are best capable
CC of withstanding a particular disease or environmental stress; (3) for
CC enhancing or inhibiting production of a biosynthetic product in a plant;
CC (4) as probes in mapping and in diagnosis, in genetic modification and
CC for screening purposes, to generate additional copies of the nucleic
CC acids, to generate ribozymes or antisense oligonucleotides, and as single
CC stranded DNA probes or as triple-strand forming oligonucleotides; and (5)
CC for generating genetically modified transgenic organisms. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC USPTO web site
XX
SQ Sequence 469 BP; 133 A; 116 C; 90 G; 130 T; 0 U; 0 Other;
Query Match 28.6%; Score 326; DB 6; Length 469;
Best Local Similarity 100.0%; Pred. No. 2.4e-96;
Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 815 TCGCAGTAATAGTCATCATCTAGCGACTCAGGCAACAACTGGAGATCCACCATGG 874
Db 469 TCGCAGTAATAGTCATCATCTAGCGACTCAGGCAACAACTGGAGATCCACCATGG 410
QY 875 TGCTGAGCAACAAAGCTCCACGGTGGCATGAGCAGTTACGTTGCTGGTCTTAAATTTCC 934
Db 409 TGCTGAGCAACAAAGCTCCACGGTGGCATGAGCAGTTACGTTGCTGGTCTTAAATTTCC 350
QY 935 ATGGTCGAGTCAAGTGGCTTCGGTTAAGAACTTTTCAGCTTGGCAGTTAGTACTGTG 994
Db 349 ATGGTCGAGTCAAGTGGCTTCGGTTAAGAACTTTTCAGCTTGGCAGTTAGTACTGTG 290
QY 995 AAGCAGGGCAGACATCTGAGAGATCATACTCCAGTTTGGGAAAGTTGGGAAGGACATGT 1054
Db 289 AAGCAGGGCAGACATCTGAGAGATCATACTCCAGTTTGGGAAAGTTGGGAAGGACATGT 230
QY 1055 TTACCATGGATTATGGATATCCGATTTCTGCGCTTTTCAACGGTTTGTCTATCTGCTGAGCA 1114

Db 229 TTACCATGATTATGGATATCCGATTTCTGCGTTTCAAGCGTTTGCTATCTGCTGAGCA 170

QY 1115 GTTTTGAACACGAAATGCTGTGAA 1140

Db 169 GTTTTGAACACGAAATGCTGTGAA 144

RESULT 14

ADW17065

ID ADW17065 standard; cDNA; 1887 BP.

AC ADW17065;

XX 24-MAR-2005 (first entry)

DT Pinus radiata transcription factor cDNA TUBBY family Seq 808.

XX gene; ss; plant; transcription; gene regulation; gene expression;

XX transgenic plant; drought resistance; disease resistance; salt tolerance;

KW cold tolerance; freezing tolerance; flowering; flavor enhancer;

KW flower color.

XX

OS Pinus radiata.

XX WO2005001050-A2.

XX 06-JAN-2005.

XX 07-JUN-2004; 2004WO-US017965.

XX 06-JUN-2003; 2003US-0476189P.

XX (ARBO-) ARBORGEN LLC.

PA Bloksberg LM, Bryant C, Connell MB, Emerson SJ, Frost MJ;

PI Forster RLS, Grigor M, Higgins C, Lasham A, Lund ST, Magnusin A;

PI Phillips J, Puthigae S, Veerakone S, Westwood C, Gause K, Wood M;

XX WPI; 2005-075542/08.

DR P-PSDB; ADW17850.

XX New polynucleotides isolated from plants encoding transcription factors,

PT and polypeptides encoded by such polynucleotides, useful for regulating

PT gene transcription and gene expression.

XX Claim 3; SEQ ID NO 808; 1265pp; English.

XX This invention relates to novel isolated plant nucleic acid molecules, or

CC variants thereof, that encode transcription factors. Specifically, it

CC refers to transcription factor proteins that are capable of binding to

CC DNA in order to regulate gene transcription and gene expression in a

CC plants, in particular *Eucalyptus grandis* and *Pinus radiata*. The present

CC invention describes DNA constructs containing DNA encoding a

CC transcription factor that regulates the promoter, which is operably

CC linked to the desired nucleic acid to be expressed. It further provides

CC transgenic plants expressing a transcription factor that confers a trait

CC to the plant such as increased drought, salt or disease tolerance, height

CC change, enhanced cold/frost tolerance, enhanced color, health and

CC nutritional characteristics, as well as improved taste, starch

CC composition, flower longevity and germination, amongst others.

CC Accordingly, such plants that are successfully transfected with a DNA

CC construct can be characterized by a difference in flower color, petal or

CC leaf shape and size, aroma or plant height. This polynucleotide is a

CC plant transcription factor cDNA sequence of the invention.

XX

SQ Sequence 1887 BP; 492 A; 397 C; 459 G; 539 T; 0 U; 0 Other;

Query Match 28.5%; Score 324.6; DB 14; Length 1887;

Best Local Similarity 59.5%; Pred. No. 1.5e-95;

Matches 656; Conservative 0; Mismatches 414; Indels 33; Gaps 5;

QY 68 CAACCGCTAATAGTTTCAAGCCCTTTTCAGCTGTGCGAGCTCCCGGAGGAGCTGCTTAGAG 127

652 CAGCCGCGATGTTTTGCAATGAGAGCTGCTGGCGGAATTTGCCCTATAGCTGCTCAGAG 711

QY 128 AAATCTGATAGGTTGAGACTGTTGACGGCGGCGATTTGGCGTCCGCGCGAAACGTGG 187

Db 712 ATATTAATTGAGAGAAATTGAGAGCAGCGATCACC---TGGCTGTCGCGAGGAAAGTTG 768

QY 188 TGGCTTGTGCGCGGTTTTGCTGCTAGCTGAGAGATTTCTCACCAAGGAGATTGTAGCTGTTT 247

Db 769 TGGTTTGTGCTGCGGTGTCAGAACTGAGGAAATTTATCAAGGAGATTGTGAACCGC 828

QY 248 CTGAATTTCTCTTAAATTTGACTTTTCCCTATCTCCCTCAAGCAGTCTGTCCTCAAGAGATT 307

Db 829 CGCTGCTGTCGCGAGTGTCTCATTCTCTGCTTGAAGCAGCTGTCGACCGAAGATT 888

QY 308 CTCATAGTTCAATGCTTTTATAAAAGTAATCGAAATCTCAATCGTATCATCTCTATCTCG 367

Db 889 CTGCTGTGCAATGTTTTTATCAAGCAAGATAGACTGTCATCTGCTTATTTATTTGTTCTG 948

QY 368 GATTAACCTACCTTTTGAACGATTAACGGAATGCGGAAAGTTTCTTCTGCTGCTTAAAGCTGAAGC 427

Db 949 GTTTGGCTTCAACGATTTGTGAAATGGAATTTCTTCTGCTGCTCGAAATTCAGAA 1008

QY 428 GCGCAACTTGCACCTGATTAATCATCTCTTTTGGCTTTCAGACGATATCTCAAGAGAGAACA 487

Db 1009 AGGCTACAAGCACTGATTAATCATCTCTGTTGATGAGAGGATATGTCAAGGGAAGCA 1068

QY 488 ACGGCTATCTTGGGAGATGAGATCGAATCTTCTTGGAAACAAATTCACGGCTTTGATG 547

Db 1069 ATACCTACGTAGGAAAGCTGAGTCAAAATTTCTTTGGAAACCAAGTTTCAACAATATATGATA 1128

QY 548 GTAGTCAGAC-----CGGAGCAGCGAAGATCGAAGAGAGCGCTCTTCTAATTTCA 598

Db 1129 GTCAACCTTCTCATAGTGTGTGTAGTCTCAAGTCTGCGAGGCGACGGGTTGGTTCAA 1188

QY 599 TCAAAAGTTTCACTAGAGTTCTCTCAGGGAAGTTTACCCACTCGCTCACAATTTTCATACGAGT 658

Db 1189 ASCAAGTTTCTCCAGAGTTTCCAGCTAGTAGCTATAATGTTGCTCACAATGCTCATGAAT 1248

QY 659 TAAACGCTTTAGGCTCTCGGGACCGAGAGAAATGCGTTGTCATCATGATCAATAACCTTA 718

Db 1249 TGAATGTTTGGGTCAAGGGGTCCAAGGAGTCAATGATGTTACCATGCTTCTTATACCAG 1308

QY 719 TGAGCATGTCGAGTCGCGAGGAGTAGTAGTTC-----AACATCCATAAGCTCTTTT 772

Db 1309 CGTCAGTCGTAGAAGGAGATTGAACATTTTGAATTTCCATCAAGTACTCTAG 1368

QY 773 CCAGTCGGTCAATCACCAGTCTTTTAGTCTCACTCAAAACCATTTGCGCAGTAATAGTGCAT 832

Db 1369 TGGAGCTCTAGCATCTTTCCGTTCTCTCGGTCGAAGTCAGTCATTAGTGAAGTAATG 1428

QY 833 CATGTAGCGACTCAGGCAAC-----AACCTGGGAGATCCACATTTGGTGTGAGCAACA 886

Db 1429 CTTTTTCTGGGCTCTCAATAGTGGGAACCTTGAAGATGATCCATTTGATCTTTGAAGAACA 1488

QY 887 AGCTCCAGGTGGCATGAGCAGTTACGTTGCTGCTTAAATTTCCATGCTCGAGTCA 946

Db 1489 AAGCCCAAGTGGCATGAGCAGTTCATGCTGCTTTGAATTTTCAAGGACGGGTGA 1548

QY 947 CAGTGGCTTCGGTTAAGAACTTTTCAAGTGTGCGCAGTTAGTG-----ACTGTGAAG 997

Db 1549 CAGTTGCATCTGTAAGAACTTTTCACTTTTGGAGCTACTAGCGCTGCACATCTGTCT 1608

QY 998 CAGGGCAGACATCTGAGAGGATCATATCTCCAGTTTGGGAAAGTTGGGAAGGACATGTTTA 1057

Db 1609 CTCAGAGTGATCAGGTTAAGGTTTATCTCCAGTTTGGAAAGGTTGGAAGGATTTTTTCA 1668

QY 1058 CCATGATTTATGGATATCCGATTTCTGCGTTTCAAGCGTTTCTGCTATCTGCTGAGCAGTT 1117

Db 1669 CTATGGATTTATCGATATCCCTGTGCTGAGCTTTTTCAGGCATTTGCCATCTGCTTAGCAGTT 1728

QY 1118 TTGAAACCAAGAAATTCCTGTGAA 1140

Db 1729 TCGACACTAAGTTGGCTGTGAA 1751

```
RESULT 15
ADW17066
ID ADW17066 standard; cDNA; 1923 BP.
XX AC
XX ADW17066;
XX DT
XX DE
XX DE
XX KW gene; ss; plant; transcription; gene regulation; gene expression;
XX KW transgenic plant; drought resistance; disease resistance; salt tolerance;
XX KW cold tolerance; freezing tolerance; flowering; flavor enhancer;
XX KW flower color.
XX OS
XX OS Pinus radiata.
XX PN
XX PN WO2005001050-A2.
XX XX
XX PD 06-JAN-2005.
XX PF 07-JUN-2004; 2004WO-US017965.
XX PR 06-JUN-2003; 2003US-0476189P.
XX XX
XX XX (ARBO-) ARBorgen LLC.
XX PI Bloksberg LN, Bryant C, Connett MB, Emerson SJ, Frost MJ;
XX PI Forster RLS, Grigor M, Higgins C, Lasham A, Lund ST, Magusin A;
XX PI Phillips J, Puthigae S, Veerakone S, Westwood C, Gause K, Wood M;
XX XX
XX DR WPI; 2005-075542/08.
XX DR P-PSDB; ADW17851.
XX XX
XX PT New polynucleotides isolated from plants encoding transcription factors,
XX PT and polypeptides encoded by such polynucleotides, useful for regulating
XX PT gene transcription and gene expression.
XX XX
XX PS Claim 3; SEQ ID NO 809; 1265pp; English.
XX CC
XX CC This invention relates to novel isolated plant nucleic acid molecules, or
XX CC variants thereof, that encode transcription factors. Specifically, it
XX CC refers to transcription factor proteins that are capable of binding to
XX CC DNA in order to regulate gene transcription and gene expression in a
XX CC plants, in particular Eucalyptus grandis and Pinus radiata. The present
XX CC invention describes DNA constructs containing DNA encoding a
XX CC transcription factor that regulates the promoter, which is operably
XX CC linked to the desired nucleic acid to be expressed. It further provides
XX CC transgenic plants expressing a transcription factor that confers a trait
XX CC to the plant such as increased drought, salt or disease tolerance, height
XX CC change, enhanced cold/ frost tolerance, enhanced color, health and
XX CC nutritional characteristics, as well as improved taste, starch
XX CC composition, flower longevity and germination, amongst others.
XX CC Accordingly, such plants that are successfully transfected with a DNA
XX CC construct can be characterized by a difference in flower color, petal or
XX CC leaf shape and size, aroma or plant height. This polynucleotide is a
XX CC plant transcription factor cDNA sequence of the invention.
XX XX
XX SQ Sequence 1923 BP; 514 A; 374 C; 457 G; 578 T; 0 U; 0 Other;
XX
Query Match 28.2%; Score 322; DB 14; Length 1923;
Best Local Similarity 59.4%; Pred. No. 1.1e-94;
Matches 641; Conservative 0; Mismatches 400; Indels 39; Gaps 4;
XX 97 TGGTCGGAGCTCCCGGAGGAGCTGCTTAGAGAAATCCTGATTAGGCTTCGAGACTGTTGAC 156
DB 545 TGGCGGAATATGCCCTCGAGTTGATAGAGATGTTATTTCAGAGAAATTGAG---GAGAGC 601
XX 157 GGCGGCGATTGGCGGTCCGGCGGAAACGTTGGTGGCTGTGCCGGCGTTTGTGTAGCTGG 216
DB 602 GAGAGCAGTTGGCTTTTCAGGAGAGAGTGTGGTGGCATGTGCTGGGGTTGTATGACATGG 661
```

```
QY 217 AGGATTTCTCACAAAGGAGATTGTAGCTGTTCTGTAATTTCTCTCTAAATTTGACTTTCCCT 276
DB 662 AGGCAAAATTACAAAGGAGCTTGTCCGATACCTGAGCTATCTGGGAAGCTTACCTTCCCC 721
QY 277 ATCTCCCTCAAGCAGTCTGTGTCGAAGAGATTCTCTAGTTCAATGCTTTTATAAAACGTAT 336
DB 722 ATTTCTTTGAAGCAGCTGGATCAAGGGAGCCCTCTGTACAGTGTTTTCATCAAGCGAGAT 781
QY 337 CGAATACTCAATCGTATCATCTCTATCTCCGATTAACCTCTTTTTCAGCGATAACGG 396
DB 782 AGGGCTACATCTACCTATCAATTTGTTCTCTGTCTGAATACACACTTTTCTGAAATGA 841
QY 397 AAGTTTCTTTCTGCTGCTTCTAAAGCTGAAGCGCGCAACTTGCACCTGATTACATCTCT 456
DB 842 AAATTTCTTTCTGCGCTCGGAAATTCAGCGGGCTACAAGTACAGATATATTATTCT 901
QY 457 TTGCGTTCAGACGATATCTCAAAGAGAAGCAACGGTATCTTTGGGAGAGATGAGATCGAAC 516
DB 902 CTTAATGCAGACGATATGTCAAGGGGTAGCAATACATATGTTGGAAATTTGAGGTCAAC 961
QY 517 TTCCTTGGAAACAAATTTCAACGGTCTTTGATGG-----TAGTCAGACCGGAGCA 564
DB 962 TTTCTTGGTACTTAAGTTTACAATATATATAGTACCCAACTCTCTCAGTGGGCAATTGCA 1021
QY 565 GCGAAGATGCAGAGAGCGGCTCTTCTAAATTTTCATCAAAAGTTTTCACCTAGAGTTCTCTCAG 624
DB 1022 TCACAAAGCGTGCAGGTGCGAGGTAGGAGCAAAACAAAGTTTCTTCCAAGAGTTCACGCT 1081
QY 625 GGAAGTTACCCCATCGCTCACAATTTTCATACGAGTTAAAGCTTTAGGCTCTCGGGAGCG 684
DB 1082 GGTAGCTATAATATTTGCTCACAATGGCTATGAATTTGAATGTCTTGGGAACTAGGGGTCCA 1141
QY 685 AGAAGATCGGTTGCATCATGATACATACCTATGAGCATGCTGGAGTCGGGAGGAGTA 744
DB 1142 AGACGATGCAATGATCATGATCTTATTCAGCTTCTGCTGTGAGCAGGAGGCGAGT 1201
QY 745 GTAGCTTCAACATCCATCAAGCTCTTTTCCAGTCCG-----TCATCACCACCA 789
DB 1202 GCCCAACTCTGTTGAACCTTCGGTCTGCTAGTCTAGACCAGTCTCTTGCACTTTGCTCT 1261
QY 790 GTCTTTAGGTCTCACTCAAAACCAATTCGCGCAGTAATAGTGCATCATGTAGCGACTCAGGC 849
DB 1262 ACTTTTGGCTCCAAAGTCTGTCATTTTAGAAAGTTTCATCTCTGAGCCACTTACCGGT 1321
QY 850 AACACCTGGGAGATCCACCATTTGCTGTGACCAACAAGCTCCACGTTGGCATGAGCAG 909
DB 1322 AGCAACTTTGAAAGATGAACCATTTGGTTTAAATAAATAAGCGCCAAAGTGGCATGAGCAG 1381
QY 910 TTACGCTGCTGCTGCTTAAATTTCCATGCTGAGTCACAGTGGCTTCGGTTAAGAACTTT 969
DB 1382 CTTCAATGTTGGTGTGTTGAAATTTCAAGGGGCTGTGACTGTGGCTTCTGTTAAGAAATTT 1441
QY 970 CAGCTTTGGCAGTTAGTACTGTCAAGCA-----GGCAGACATCTCAGAGGATC 1020
DB 1442 CAGCTTTGTCAGCACCTTGAGCCTAATCCATCGGTTTCTCAGAATGATTGTGATAAGGTC 1501
QY 1021 ATACTCCAGTTTGGGAAAGTTGGGAAGACATGTTTACCATGATTTATCGATTCGATTCGATT 1080
DB 1502 ATACTCCAATTTGGAAAGGTTGGAAAGGACATGTTTCAACAATGATATCGCTACCCGCTC 1561
QY 1081 TCTGGGTTTCAAGCGTTTGGCTATCTGCCTGAGCAGTTTGGAAACCAAGAAATTCGCTGTGAA 1140
DB 1562 TCAGCTTTTCAAGCTTTTGGCCATATAGCTTAGTAGTTTGTGACACAAACTTGTGTTGAA 1621
```

Search completed: December 26, 2005, 03:41:22
Job time : 803 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 26, 2005, 01:47:21 ; Search time 6049 Seconds

(without alignments)
10712.768 Million cell updates/sec

Title: US-10-763-042-20

Perfect score: 1140

Sequence: 1 atgacgttcgaagttaact.....aaaccagaattgcctgtgaa 1140

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.in.*

3: gb.env.*

4: gb.cm.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pr.*

9: gb.ro.*

10: gb.sts.*

11: gb.sv.*

12: gb.un.*

13: gb.vi.*

14: gb.htg.*

15: gb.pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1140	100.0	1143	6	CQ804532 Sequence
2	1140	100.0	1143	15	AF487270 Arabidops
3	1037.4	91.0	1595	15	BT004092 Arabidops
C	598	52.5	86022	15	ATAC011623 Arabidops
	572.8	50.2	1170	15	BT020601 Arabidops
6	572.8	50.2	1259	15	BT010768 Arabidops
7	567.8	49.8	1143	15	AY046922 Arabidops
8	350.6	30.8	1781	15	BT013601 Lycopersi
9	342.6	30.1	1221	6	CQ804888 Sequence
10	342.6	30.1	1221	15	AY045774 Arabidops
11	342.6	30.1	1252	15	AY096604 Arabidops
12	342.6	30.1	1943	15	AY074273 Arabidops
13	336.2	29.5	1687	15	AK071159 Oryza sat
14	336.2	29.5	1948	15	AK060587 Oryza sat
15	262.8	23.1	2102	15	AK102370 Oryza sat
16	250.4	22.0	1185	15	AY045773 Arabidops
17	248.8	21.8	1689	15	BT008627 Arabidops
18	240.4	21.1	1571	15	AK062089 Oryza sat

19	240.4	21.1	2345	15	AK064855	Oryza sat
20	210.2	18.4	1167	15	AF487268	Arabidops
21	210.2	17.6	1722	15	CAR400860	Cicer ari
22	195.4	17.1	1338	15	AF487271	Arabidops
23	193	16.9	1140	15	AY092403	Arabidops
24	193	16.9	1171	15	AY059088	Arabidops
25	193	16.9	1388	15	AF370146	Arabidops
26	187.4	16.4	85746	15	AC051627	Genomic S
27	184.2	16.2	878	15	AF386521	Pyrus com
28	184	16.1	1290	15	AY046921	Arabidops
29	184	16.1	1321	15	AY122972	Arabidops
30	184	16.1	1567	15	AF424565	Arabidops
31	184	16.1	1582	15	AY072154	Arabidops
32	179.8	15.8	1539	15	AB023895	Lemna pau
33	176	15.4	1341	15	AK102298	Oryza sat
34	171.2	15.0	2027	15	AK103583	Oryza sat
35	170.2	14.9	1103	15	BT018325	Zea mays
36	169.6	14.9	2023	15	AK069622	Oryza sat
37	169.2	14.8	1257	15	AK106164	Oryza sat
38	165	14.5	1630	15	AK098851	Oryza sat
39	165	14.5	1685	15	AK061747	Oryza sat
40	163.4	14.3	2129	15	AK100078	Oryza sat
41	163.4	14.3	3119	15	AK120706	Oryza sat
42	160.8	14.1	1824	15	AK066910	Oryza sat
43	160.4	14.1	1994	15	AK102221	Oryza sat
44	159.2	14.0	1800	15	AK104333	Oryza sat
45	159.2	14.0	1826	15	AK098928	Oryza sat

ALIGNMENTS

RESULT 1					
CQ804532	Sequence 943 from Patent WO2004035798.	1143 bp	DNA	linear	PAT 10-MAY-2004
LOCUS					
DEFINITION					
ACCESSION					
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					

Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi

REFERENCE 1
AUTHORS Inze, D., de Veylder, L. and Vlieghe, K.
TITLE Identification of novel e2f target genes and use thereof
JOURNAL Patent: WO 2004035798-A 943 29-APR-2004;
CropDesign N.V. (BE)

FEATURES
source 1. .1143
Location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN

Query Match	100.0%;	Score 1140;	DB 6;	Length 1143;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1140;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ATGACGTTCCGAAGTTTACTCCAGGAAATGCGGTCTAGGCCACACCGTGTAGTTTCAGGCC	60	
Db	1	ATGACGTTCCGAAGTTTACTCCAGGAAATGCGGTCTAGGCCACACCGTGTAGTTTCAGGCC	60	
Qy	61	GCGGCTCAACCGCTAATAGTTTCAGACCTTTTCAGCTGTTCGAGCTCCCGGAGGAGCTG	120	
Db	61	GCGGCTCAACCGCTAATAGTTTCAGACCTTTTCAGCTGTTCGAGCTCCCGGAGGAGCTG	120	
Qy	121	CTTAGAGAAATCCTGATTAGGGTTGAGACTGTTGACGCGGCGGATTTGGCGCTCGCGCGA	180	
Db	121	CTTAGAGAAATCCTGATTAGGGTTGAGACTGTTGACGCGGCGGATTTGGCGCTCGCGCGA	180	
Qy	181	AACGTGGTGGCTTGTGCGCGGCTTTGTCTGTAGCTGGAGGATTTCTACCAAGGAGATTGTA	240	

Db 181 AACGTTGGCTTGTGCGCGGTTTCTCGTAGCTGGAGATTCTCAACGAGGATTGTA 240
Qy 241 GCTGTTCCCTGAATTCCTCTAAATGACTTCCCTATCTCCCTCAAGCAGTCTGGTCCA 300
Db 241 GCTGTTCCCTGAATTCCTCTAAATGACTTCCCTATCTCCCTCAAGCAGTCTGGTCCA 300
Qy 301 AGAGATTCTCTAGTTCAATGCTTTATAAAGCGTAATCGAAATCTCAATCGTATCATCTC 360
Db 301 AGAGATTCTCTAGTTCAATGCTTTATAAAGCGTAATCGAAATCTCAATCGTATCATCTC 360
Qy 361 TATCTCGGATTAACTACCTCTTTGACGGGATAACGGGAAGTTTCTTCTGCTGCTCTTAAG 420
Db 361 TATCTCGGATTAACTACCTCTTTGACGGGATAACGGGAAGTTTCTTCTGCTGCTCTTAAG 420
Qy 421 CTGAAGCGGCGCAACTTGGCATGANTACATCATCTCTTTGGGTTTCAGAGATATCTCAAG 480
Db 421 CTGAAGCGGCGCAACTTGGCATGANTACATCATCTCTTTGGGTTTCAGAGATATCTCAAG 480
Qy 481 AGAAGCAACGCGTATCTTGGGGAATGAGATCGAATCTCTTGGACAAATTCACGGTC 540
Db 481 AGAAGCAACGCGTATCTTGGGGAATGAGATCGAATCTCTTGGACAAATTCACGGTC 540
Qy 541 TTTGATGTAGTCAGACCGGAGCAGCAAGATGCAAGAGCGGCTCTTCTTAATTTTCATC 600
Db 541 TTTGATGTAGTCAGACCGGAGCAGCAAGATGCAAGAGCGGCTCTTCTTAATTTTCATC 600
Qy 601 AAAGTTTACCTAGAGTTCTCAGGGAAGTTACCCCATCGCTCACATTTTCAACAGATTA 660
Db 601 AAAGTTTACCTAGAGTTCTCAGGGAAGTTACCCCATCGCTCACATTTTCAACAGATTA 660
Qy 661 AAGCTTTAGGCTCTCGGGACCGAGAAATGCGTTGTCATCATGATCAATACCTATG 720
Db 661 AAGCTTTAGGCTCTCGGGACCGAGAAATGCGTTGTCATCATGATCAATACCTATG 720
Qy 721 AGCATCGTGGAGTCGCGAGAGTAGTCTTCAACATCAATAGCTCTTTTCCAGTCGG 780
Db 721 AGCATCGTGGAGTCGCGAGAGTAGTCTTCAACATCAATAGCTCTTTTCCAGTCGG 780
Qy 781 TCATCACCATGTTTGGTCTCTACTCAAAACCATTCGCGAGTAATAGTCATCATAGC 840
Db 781 TCATCACCATGTTTGGTCTCTACTCAAAACCATTCGCGAGTAATAGTCATCATAGC 840
Qy 841 GACTCAGGCAACACTTGGGAGATCCACCATTTGGTGTGAGCAACAAAGCTCCACGGTCG 900
Db 841 GACTCAGGCAACACTTGGGAGATCCACCATTTGGTGTGAGCAACAAAGCTCCACGGTCG 900
Qy 901 CATGACGAGTTAGCTTGTGCTTAAATTTCCATGTCGAGTCACAGTGGCTTCGGTT 960
Db 901 CATGACGAGTTAGCTTGTGCTTAAATTTCCATGTCGAGTCACAGTGGCTTCGGTT 960
Qy 961 AAGAACTTTTCCAGCTTGTGGCAGTTAGTCACTGTGAAGCAGGCGAGACATCTGAGAGATC 1020
Db 961 AAGAACTTTTCCAGCTTGTGGCAGTTAGTCACTGTGAAGCAGGCGAGACATCTGAGAGATC 1020
Qy 1021 ATACTCCAGTTTGGAAAGTTGGGAGGACATGTTTACCATGATATGATATCCGATT 1080
Db 1021 ATACTCCAGTTTGGAAAGTTGGGAGGACATGTTTACCATGATATGATATCCGATT 1080
Qy 1081 TCTGCGTTTCAAGCGTTTGTCTATCTGCTGAGCAGTTTGTGAACACAGAAATGCTGTGAA 1140
Db 1081 TCTGCGTTTCAAGCGTTTGTCTATCTGCTGAGCAGTTTGTGAACACAGAAATGCTGTGAA 1140

RESULT 2
AF487270
LOCUS AF487270
DEFINITION Arabidopsis thaliana tubby-like protein TULP9 (TULP9) mRNA,
complete cds.
ACCESSION AF487270
VERSION AF487270.1 GI:33320970
KEYWORDS Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi
1 (bases 1 to 1143)
Lai,C.P., Lee,C.L., Chen,P.H., Wu,S.H., Yang,C.C. and Shaw,J.F.
Molecular Analyses of the Arabidopsis TUBBY-Like Protein Gene
Family
Plant Physiol. 134 (4), 1586-1597 (2004)
15064372
2 (bases 1 to 1143)
Lai,C.P. and Shaw,J.F.
Cloning and characterization of cDNAs for tubby-like protein 9
Unpublished
3 (bases 1 to 1143)
Lai,C.P. and Shaw,J.F.
Direct Submission
Submitted (25-FEB-2002) Institute of Botany, Academia Sinica,
Nankang, Taipei, Taiwan 11529, Republic of China
Location/Qualifiers
1..1143
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
1..1143
/gene="TULP9"
1..1143
/gene="TULP9"
/note="putative tub family protein"
/codon_start=1
/product="tubby-like protein TULP9"
/protein_id="AA06243.1"
/db_xref="GI:33320971"
/translation="MTFRSLIQEMRSRPHRVVHAAASTANSDDPFSWSLPELLREI
LIRVETVDGDMPSRNVACQVCSWRILTKETIVAFEFSSKLTFFISLQSGPRD
SLVQCFIKRNRNTQSHLYLGLTSLTDNGKFLAASKLKRATCTDYIISLSDDISK
RSNAVLYGRNNTFLGTFTFDGSGTAAKMKSRSNPFIKVSPPVQSGSYPIAHISY
ELNVLSGRSPRMRICIMDTIPMSIYESRGVASTISISFSSRSSPVRSHSKPLRNS
ASCSDGNLGDPLVLSNKAPRWEQRLCWLNFHGRVTVAASKNFQLVAVSDCEAG
QTSERIILQPGKVGKDMFTMDYGYPISAFOAFICLSSFETRICE"

ORIGIN

Query Match 100.0%; Score 1140; DB 15; Length 1143;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGACGTTCCGAGTTTACTCCAGGAAATGCGGTCTTAGGCCACACCGTGTAGTTACGCC 60
Db 1 ATGACGTTCCGAGTTTACTCCAGGAAATGCGGTCTTAGGCCACACCGTGTAGTTACGCC 60
Qy 61 GCGGCTCAACCGCTAATAGTTTCAGACCTTTTCAGTGGTTCGAGGCTCCGGAGGAGCTG 120
Db 61 GCGGCTCAACCGCTAATAGTTTCAGACCTTTTCAGTGGTTCGAGGCTCCGGAGGAGCTG 120
Qy 121 CTTAGAGAAATCCTGATTAGGGTTGAGACTGTTCAGCGCGGCGATTGGCGCTCGCGCGGA 180
Db 121 CTTAGAGAAATCCTGATTAGGGTTGAGACTGTTCAGCGCGGCGATTGGCGCTCGCGCGGA 180
Qy 181 AACGTTGGTGGCTTGTCCCGCGGTTTTGTCTAGCTGGAGGATTTCTACCAAGGAGATTGA 240
Db 181 AACGTTGGTGGCTTGTCCCGCGGTTTTGTCTAGCTGGAGGATTTCTACCAAGGAGATTGA 240
Qy 241 GCTGTTCTGGAATTCCTCTTAATTTGACTTTTCCCTATCTCCCTCAAGCAGTCTGTGCA 300
Db 241 GCTGTTCTGGAATTCCTCTTAATTTGACTTTTCCCTATCTCCCTCAAGCAGTCTGTGCA 300
Qy 301 AGAGATTCTCTAGTTCAATGCTTTTATAAAGCGTAATCGAAATCTCAATCGTATCATCTC 360
Db 301 AGAGATTCTCTAGTTCAATGCTTTTATAAAGCGTAATCGAAATCTCAATCGTATCATCTC 360
Qy 361 TATCTCGGATTAACTACCTCTTTTGAACGAGTAAACGGGAAGTTTCTTCTGCTGCTTAAG 420
Db 361 TATCTCGGATTAACTACCTCTTTTGAACGAGTAAACGGGAAGTTTCTTCTGCTGCTTAAG 420

Db 327 AACGTGGTGGCTTGTGCGCGCTTTGTCTAGCTGGAGGATTTCTCACCAGGAGATTGTA 386
QY 241 GCTGTTCTCGAATTTCTCTCTAAATGACTTTCCCTATCTCCCTCAAGCAGTCTGGTCA 300
Db 387 GCTGTTCTCGAATTTCTCTCTAAATGACTTTCCCTATCTCCCTCAAGCAGTCTGGTCA 446
QY 301 AGAGATTCTCTAGTTCAATGCTTTTATAAACGTAATCGAAATACTCAATCGTATCATCTC 360
Db 447 AGAGATTCTCTAGTTCAATGCTTTTATAAACGTAATCGAAATACTCAATCGTATCATCTC 506
QY 361 TATCTCGGATTAACTACCTCTTTGACGGATAACGGGAAGTTTCTTCTGCTGCTTCTTAAG 420
Db 507 TATCTCGGATTAACTACCTCTTTGACGGATAACGGGAAGTTTCTTCTGCTGCTTCTTAAG 566
QY 421 CTGAAGCGCGCAACTGCACTGATTAATCATCATCTCTTTCGGTTTCAGCATATCTCAAG 480
Db 567 CTGAAGCGCGCAACTGCACTGATTAATCATCATCTCTTTCGGTTTCAGCATATCTCAAG 626
QY 481 AGAAGCAACGCGTATCTTGGGAGATGAGATCGAACTTCTTGGAAACAAAATTCACGGTC 540
Db 627 AGAAGCAACGCGTATCTTGGGAGATGAGATCGAACTTCTTGGAAACAAAATTCACGGTC 586
QY 541 TTTGATGTPAGTCAGACCGGAGCAGCAAGATGCAAGACGCGCTCTTCTTAATTTTCATC 600
Db 687 TTTGATGTPAGTCAGACCGGAGCAGCAAGATGCAAGACGCGCTCTTCTTAATTTTCATC 746
QY 601 AAAGTTTACCTAGATTCTCTCAGGGAAGTTTACCCATCGCTCACATTTTCATACAGTTTA 660
Db 747 AAAGTTTACCTAGATTCTCTCAGGGAAGTTTACCCATCGCTCACATTTTCATACAGTTTA 806
QY 661 AACGCTTTAGGCTCTC----- 676
Db 807 AACGCTTTAGGCTCTCGGTAACAGTTTTCATTCAGTATGATTAATGCTTCTTCTAGT 866
QY 677 -----GGGACCGGAGAG 689
Db 867 CTTTGGACCATCTGATTTTAAACAGTCTTTTGAGGGTATTAAATAACAGGGGACCGAGAAG 926
QY 690 AATGCGTTGCATGATGATCAATACCTATGAGCATCTGTTGAGTCGCGAGGAGTAGTAGC 749
Db 927 AATGCGTTGCATGATGATCAATACCTATGAGCATCTGTTGAGTCGCGAGGAGTAGTAGC 986
QY 750 TTCACATCCATAAGCTCTTTTCCAGTCGGTCAATCACCAGTCTTTTGTCTTCACTCAAA 809
Db 987 TTCACATCCATAAGCTCTTTTCCAGTCGGTCAATCACCAGTCTTTTGTCTTCACTCAAA 1046
QY 810 ACCATTGCGCAGTAATAGTGATCATGTAGCGACTCAGGCAACAACTGGGAGATCCACC 869
Db 1047 ACCATTGCGCAGTAATAGTGATCATGTAGCGACTCAGGCAACAACTGGGAGATCCACC 1106
QY 870 ATTGTTGCTGAGCAACAAAGCTCCAGGTGGCATGAGCATGATGAGTGTGCTGCTTAA 929
Db 1107 ATTGTTGCTGAGCAACAAAGCTCCAGGTGGCATGAGCATGATGAGTGTGCTGCTTAA 1166
QY 930 TTTTCCATGTCGAGTCACAGTGGCTTTCGGTTAGAACTTTTCAGTTTGTGCGAGTTAGTGA 989
Db 1167 TTTTCCATGTCGAGTCACAGTGGCTTTCGGTTAGAACTTTTCAGTTTGTGCGAGTTAGTGA 1226
QY 990 CTGTGAAGCAGGGCAGACATCTGAGAGGATCATACTCCAGTTTGGGAAAGTTGGGAAGGA 1049
Db 1227 CTGTGAAGCAGGGCAGACATCTGAGAGGATCATACTCCAGTTTGGGAAAGTTGGGAAGGA 1286
QY 1050 CATGTTTACATGGAATATGATATCCGATTTCTGCGTTTCAAGCGTTTGTCTATCTGCCT 1109
Db 1287 CATGTTTACATGGAATATGATATCCGATTTCTGCGTTTCAAGCGTTTGTCTATCTGCCT 1346
QY 1110 GAGCAGTTTGAACCAAGATTCGCTGTGAA 1140
Db 1347 GAGCAGTTTGAACCAAGATTCGCTGTGAA 1377

RESULT 4

ATAC011623/c

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
AUTHORS
TITLE
JOURNAL
AUTHORS
TITLE
JOURNAL
COMMENT

ATAC011623 86022 bp DNA linear PLN 30-OCT-2002
Arabidopsis thaliana chromosome III BAC F24P17 genomic sequence,
complete sequence.
AC011623
AC011623.5 GI:12408729
HTG.
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 86022)
Lin,X., Kaul,S., Town,C.D., Benito,M.-I., Creasy,T.H., Haas,B.,
Rowman,C.M., Koo,H., Fujii,C.Y., Utterback,T.R., Barnstead,M.E.,
Arabidopsis thaliana chromosome III BAC F24P17 genomic sequence
Unpublished
2 (bases 1 to 86022)
Lin,X. and Kaul,S.
Direct Submission
Submitted (08-OCT-1999) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA, xlin@tigr.org
3 (bases 1 to 86022)
Lin,X.
Direct Submission
Submitted (24-JAN-2001) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA
On Jan 24, 2001 this sequence version replaced gi:12280803.
Address all correspondence to:
Xiaoying Lin
The Institute for Genomic Research
9712 Medical Center Dr.
Rockville, MD 20850, USA
e-mail: xlin@tigr.org
BAC clone F24P17 is from Arabidopsis chromosome III and is near the
molecular marker mi403.
The orientation of the sequence is from SP6 to T7 end of the BAC
clone.

Genes were identified by a combination of three methods: Gene
prediction programs including GRL (available by anonymous ftp
from arthur.epm.ornl.gov), Genefinder (Phil Green, University of
Washington), Genscan (Chris Burge,
http://www.cbs.dtu.dk/netgene/cbsnetgene.html), and NetPlantGene
(http://www.cbs.dtu.dk/netgene/cbsnetgene.html), searches of the
complete sequence against a peptide database and the Arabidopsis
EST database at TIGR (http://www.tigr.org/tdb/at/at.html).
Annotated genes are named to indicate the level of evidence for
their annotation. Genes with similarity to other proteins are named
after the database hits. Genes without significant peptide
similarity but with EST similarity are named as 'unknown' proteins.
Genes without protein or EST similarity, that are predicted by more
than two gene prediction programs over most of their length are
annotated as 'hypothetical' proteins. Genes encoding tRNAs are
predicted by tRNAscan-SE (Sean Eddy,
http://genome.wustl.edu/eddy/tRNAscan-SE/). Simple repeats are
identified by RepeatMasker (Arian Smit,
http://ftp.genome.washington.edu/RM/RepeatMasker.html). Regions of
genomic sequence that are not annotated as genes but have predicted
exons by GRL are annotated as misc features.

FEATURES
source

1. 86022
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/chromosome="III"
/map="mi403"
/clone="F24P17"
/ecotype="Columbia"
~283. 1822
/gene="F24P17.1"
/note="similar to pyruvate dehydrogenase kinase
GB:AAC97601, [Arabidopsis thaliana]"
Join(<283. 529,638. 818,916. 1063,1101. 1142,1160. 1376,

gene

mRNA

```

1511..1822)
/genes="F24P17.1"
/join(<283..529,638..818,916..1063,1101..1142,1160..1376,
1511..15169)
/genes="F24P17.1"
/codon_start=1
/product="putative pyruvate dehydrogenase kinase, 5'
partial"
/protein_id="AAF08568.1"
/db_xref="GI:6437541"
/translation="VRDWTLESFRDMRAPEIKDSGDEKDFQMIKAVKVRHNVPVM
MALGVNOLKGNNGNLDIHOFLDRFYLGRIMLIGOHVELHNPPLHTVGVIIH
TKMSPEVARNASEDARSICREYGSAPINLYGDPSEFPVYVTHLHMYELVKNNS
LRAVQRFYDRVRAPPIRIIIVADGIEDVTIKPPRSLHRFPPIIVVDEGGGIARSQ
LRFIFLYSTARNPLERDVDLGIADVPVTMAGYGYGLPIRLIYARFYFGDLQIISME
GGTDAYLHSLRIGSQBELP"
/complement(2022..2051)
/rpt_family="POLY A"
/complement(2053..2372)
/notes="exon predicted by xgrail, quality
excellent_shadowexon"
2057..2160
/misc_feature
/notes="exon predicted by xgrail, quality
excellent"
2287..2363
/notes="exon predicted by xgrail, quality excellent"
<2630..>56599
/genes="F24P17.2"
/notes="similar to RNA helicase (DRH1) GB:BAA28347
[Arabidopsis thaliana]; Pfam HMM hits: WW/rsp5/WWP domain
containing proteins, DEAD and DEAD box helicases,
Helicases conserved C-terminal domain"
/join(<2630..3712,3813..4007,4256..4324,4406..4675,
4767..5117,5215..5492,5579..56599)
/genes="F24P17.2"
/join(2630..3712,3813..4007,4256..4324,4406..4675,
4767..5117,5215..5492,5579..56599)
/genes="F24P17.2"
/codon_start=1
/product="putative RNA helicase"
/protein_id="AAF08584.1"
/db_xref="GI:6437557"
/translation="WATTEDTASAGPRVAPREDPTLPQWKGIDSGTGLYYWNPET
NTQVERPAPPHSATTTKLAIQIPVSSGQHQHQHQAPVGHVSQHQGPQPOQ
FSPQVRPQHPQHPQHPQHPQHPQHPQHPQHPQHPQHPQHPQHPQHPQHPQHPQ
MQNQSHPQHPQHPQHPQHPQHPQHPQHPQHPQHPQHPQHPQHPQHPQHPQHPQ
SLAPREGDEFGGKGTGFSQPHLPNSRSPQNTHEANAASQKTNLANLAKQKNG
POANAATVQFQPGANLHQOLGPRAPQMDOTMLHOKSHVSPFQSNNTYNNLQSRP
GNDSYNMEVPEVRGAQPLHPAMPKDIRISGPPETNADPAMGOTGCTYGHGAPAF
PKSLVRPHFVSPDVPHLSPVEIYRKHQHEVTTGENIPAPYITFESSGLPEILREL
LSAGPSPPTIQAQTWPIALQSRDIIVAIAKTSGKTLGYLIPAFILLRHRNDRNGP
TVLIIAPTRELATQIDELARFSRSRISCTCLYGAPGQPKELERGIADIVATPG
RLNDILEMWDIDFQVSLVLDLADRLMDGFEPQIRKIIVNEIPPRQTLMTATWPK
EYRKASDILVNPVQVNICRVDELAAKAI TOYVEVVPQWEXERLEQILRSQERGSK
VLI FCTSKRLCHLARSVGRHGVAVIHGDKTQGERDWTNLQFRSKSCVLIATDVA
GLDIDIRVINDIPPTVEDIVHRIKTRGATGVAPTFTEQDMWYAPDLIKVL
EGANQQVPPVQIAMRGGGGGPGYSQDRGMVNRFPDGGGTWDSGGGGRGGG
FSRGGGFGGGRGGGGRGGGGRGGGGRGGGGRGGGGRGGGGRGGGGRGGGGRGGG
RGCFGRGNMNSRSGRGGGGRGGGGRGGGGRGGGGRGGGGRGGGGRGGGGRGGG
RSRSRSGSYRSRSRSGRSGRSGRSGRSGRSGRSGRSGRSGRSGRSGRSGRSGR
VSGFDIKPPVESVNLDMNAAAENIVPTSLSERQNGVGVSEVAALVRPVUDEEP
"
5843..5962
/rpt_family="(TGG)n"
/complement(7077..>8090)
/genes="F24P17.3"
/notes="similar to hypothetical proteins GB:CAB16818,
GB:AAD5277"
/complement(7077..>8090)
/genes="F24P17.3"
/complement(7254..8090)
/genes="F24P17.3"
/notes="unknown protein"
/codon_start=1

```

```

/protein_id="AAF08569.1"
/db_xref="GI:6437542"
/translation="MASIYSSTYMLVNHPIYISNFTWIEGETLGSVFFVSVVSVYL
SATFLRSALDSPSLPFLFAQVFLKSLILCLSLVMAVGCITLITSSDSDPMA
RFLHAICFPVDKPNGLPFLFAQVFLKSLILCLSLVMAVGCITLITSSDSDPMA
VVMVGLMLRTRQSMFPFIALVTNSTMVIMYGYFLCAVGRPKRLVTDCCIQVQFV
RFGSLGWMRLREHLFGSGCTGIWGCFAAFNASLLALFNFHFSKNYVKKPTREDGKK
SD"
8262..8362
/notes="exon predicted by xgrail, quality
marginal_shadowexon"
/complement(8375..8450)
/notes="exon predicted by xgrail, quality marginal"
/genes="F24P17.4"
/notes="similar to hypothetical proteins GB:CAB16818,
GB:AAD5277"
/complement(<8719..>9615)
/genes="F24P17.4"
/complement(8719..9615)
/genes="F24P17.4"
/notes="unknown protein"
/codon_start=1
/protein_id="AAF08570.1"
/db_xref="GI:6437543"
/translation="MASVYSTLYMLVHHPIYIANFTWIEGETLGSVFFVSVVSVYL
SATFLRYVDSLPTGPRILKPIITAVHSLIILFLSLTMAVGCITLITSSDPRKLF
DAVCFPLDVKPGPLFLFAQVFLKSLILCLSLVMAVGCITLITSSDPRKLF
LCYLMLRTRQSMFPFIALVTNSTMVIMYGYFLCAVGRPKRLVTDCCIQVQFV
GLGAAMLPEHYFGSGCAGITWTVFNGVTASLLALFYNFHSKNYVKKPTREDGKK
FIFIHGESWANKAITLFSKKND"
9798..9853
/notes="exon predicted by xgrail, quality good_shadowexon"
/complement(10146..11457)
/genes="F24P17.5"
/notes="similar to putative protein GB:CAB45783"
/pseudo
/complement(12632..12715)
/notes="exon predicted by xgrail, quality
marginal_shadowexon"
<12993..>16211
/genes="F24P17.6"
/notes="similar to putative Anion exchanger family members:
GB:AAD3673, GB:AAD5295 [Arabidopsis thaliana]"
/join(<12993..13102,13175..13257,13357..13549,13667..13805,
13900..14065,14182..14339,14418..14612,14700..14795,
14879..15013,15093..15400,15493..15763,15855..>16211)
/genes="F24P17.6"
/join(12993..13102,13175..13257,13357..13549,13667..13805,
13900..14065,14182..14339,14418..14612,14700..14795,
14879..15013,15093..15400,15493..15763,15855..16211)
/genes="F24P17.6"
/notes="unknown protein"
/codon_start=1

```

Query Match
Best Local Similarity 52.5%; Score 598; DB 15; Length 86022;
Matches 1140; Conservative 0; Pred. No. 2.4e-182;

Matches 1140; Conservative 0; Indels 502; Gaps 4;

QY 1 ATGACGTTCCGAGTTTACTCCAGGAATGCGGTCTAGCCACACCGTGTAGTTACGCC 60
DB 55912 ATGACGTTCCGAGTTTACTCCAGGAATGCGGTCTAGCCACACCGTGTAGTTACGCC 55853
QY 61 GCGGCTCAACCGCTAAATAGTTTACAGCCCTTTTCAGTGTGCGAGCTCCCGGAGGAGCTG 120
DB 55852 GCGGCTCAACCGCTAAATAGTTTACAGCCCTTTTCAGTGTGCGAGCTCCCGGAGGAGCTG 55793
QY 121 CTTAGAGAAATCCTGATTAGGTTGAGACTGTTGACGGCGGCGGATTTGGCGCTCGCGCGGA 180
DB 55792 CTTAGAGAAATCCTGATTAGGTTGAGACTGTTGACGGCGGCGGATTTGGCGCTCGCGCGGA 55733
QY 181 AACGTGGTGGCTTGTGCGCGCGTGTTCGTAGTGTGAGGATTTCTACCAAGGAGATTGTA 240

Db	55732	AA	CGTGGTGGCTTGTGTGCCGGCTTTTGTGCGTAGCTGGAGGATTTCTCACCAAGGAGATTTGTA	55673
Qy	241	GC	TTCTCCTGAAATTTCTCTCTAAATTTGACTTTTCCCTATCTCCCTCAA	287
Db	55672	GC	TTTCTCCTGAAATTTCTCTCTAAATTTGACTTTTCCCTATCTCCCTCAAGCAGGTTCCTGT	55613
Qy	288	-----	-----	287
Db	55612	AA	ATTTTGTATTTTGTAGTATAAAGTGATCGAATTGAACAATAAGTTAGTCTCTTTACTT	55553
Qy	288	-----	-----	287
Db	55552	CA	TCTATAACATATGACAAATTCAGTAGCGTTTGGCGTCGCGTTTAAAGTGTCTTTGTC	55493
Qy	288	-----	-----	287
Db	55492	TC	AGTCAACAAAGCTTTTAGTCTCTTTGATCATTTGATATATAGATAAAACAGTTTCTTTG	55433
Qy	288	-----	-----GCAGTCTGGTCCAAGAGA	305
Db	55432	GT	TTTGTAAAGCTGAAGTATTTGGTACTTTTGAATGTATTATATGCAAGTCTGGTCCAAGAGA	55373
Qy	306	TT	CTCTAGTTCAATGCTTTTATAAAAAGTAATCGAAATCTCAATCGTATCATCTCTATCT	365
Db	55372	TT	CTCTAGTTCAATGCTTTTATAAAAAGTAATCGAAATCTCAATCGTATCATCTCTATCT	55313
Qy	366	CG	GAATTAACCTACCT	379
Db	55312	CG	GAATTAACCTACCTGTAAAGTCTCTGTCTTCTCTAGATTTATCAAGAAAGAGATTA	55253
Qy	380	-----	-----CTT	382
Db	55252	CT	CGTTTCTGAGTTCGAAATTTGTTAAACAGAGTCTCTTGCTAAATTTTCTGCACAGCTT	55193
Qy	383	TG	ACGATAACGGGAAGTTTCTTCTGCTGCTTCTTAAGCTGAAGCGCGCACTTCGACATG	442
Db	55192	TG	ACGATAACGGGAAGTTTCTTCTGCTGCTTCTTAAGCTGAAGCGCGCACTTCGACATG	55133
Qy	443	AT	TACATCATCTCTTTCCGTTCTCAGACGATATCTCAAGAGACGACGCGTATCTTTGGA	502
Db	55132	AT	TACATCATCTCTTTCCGTTCTCAGACGATATCTCAAGAGAGACGACGCGTATCTTTGGA	5073
Qy	503	GA	ATG	507
Db	55072	GA	TGAGGTTTGCATATAAAGTTATCTTCTTCGATATATAAGTATGCATGTTTCCAATT	55013
Qy	508	-----	-----AGATCGAACTTCTTTGGAAACAAATTTCAAGTCTTTGATGGT	549
Db	55012	TT	CTTATTTGTTTATTTGCAAGATCGAACTTCTTTGGAAACAAATTTCAAGTCTTTGATGGT	54953
Qy	550	AG	TACACCGGAGCAGCGAGATGCAAGACGCGCTCTCTTAATTTTCATCAAGTTTCA	609
Db	54952	AG	TACACCGGAGCAGCGAGATGCAAGACGCGCTCTCTTAATTTTCATCAAGTTTCA	54893
Qy	610	CT	TAGAGTTTCTCAGGGAAGTTTACCCCATCGCTCACATTTTCATACGAGTTAAACGTTTA	669
Db	54892	CT	TAGAGTTTCTCAGGGAAGTTTACCCCATCGCTCACATTTTCATACGAGTTAAACGTTTA	54833
Qy	670	GG	CTCTC	676
Db	54832	GG	CTCTCGGTAAACAGTTTTTTACATTTCAAGTATGATAATGCTTCTTCTTAGTCTTTTGACC	54773
Qy	677	-----	-----GGGGACCGAGAGAGATCGGTG	698
Db	54772	AT	TCTGATTTTAAACAGTCTTTTGTAGGGTATTAATAACAGGGGACCGAGAGAGATCGGTG	54713
Qy	699	CA	TATGATACAATACCTATGACATCTGTGGAGTCGCGAGGAGTAGTAGTCTTCAACATC	758
Db	54712	CA	TATGATACAATACCTATGACATCTGTGGAGTCGCGAGGAGTAGTAGTCTTCAACATC	54653
Qy	759	CA	TAGCTCTTTTTCCAGTCCGCTCATCAACAGTCTTTTAGTCTCTCACTCAAAAACCAATGCG	818
Db	54652	CA	TAGCTCTTTTTCCAGTCCGCTCATCAACAGTCTTTTAGTCTCTCACTCAAAAACCAATGCG	54593

Qy	819	CAGTAATAGTGCATCATGTAGCGACTCAGGCACAACTGGGAGNATCCACCATTTGGTGCT	878
Db	54592	CAGTAATAGTGCATCATGTAGCGACTCAGGCACAACTGGGAGNATCCACCATTTGGTGCT	54533
Qy	879	GAGCAACAAAGCTCCACGGTGGCATGAGCAGTTTACGTTTCTGGTGCTTAAATTTCCATGG	938
Db	54532	GAGCAACAAAGCTCCACGGTGGCATGAGCAGTTTACGTTTCTGGTGCTTAAATTTCCATGG	54473
Qy	939	TCGAGTCAACAGTGGCTTCGGTTTAAGAACTTCCAGCTTGTGGCAGTTAGTGAATGTAAGC	998
Db	54472	TCGAGTCAACAGTGGCTTCGGTTTAAGAACTTCCAGCTTGTGGCAGTTAGTGAATGTAAGC	54413
Qy	999	AGGGCAGACATCTGAGAGCATCATACTCCAGTTTGGGAAAGTTGGGAGGACATGTTTAC	1058
Db	54412	AGGGCAGACATCTGAGAGCATCATACTCCAGTTTGGGAAAGTTGGGAGGACATGTTTAC	54353
Qy	1059	CATGGATTAATGGATATCCGATTTCTGCGTTTCAACGCGTTTGCTATCTGCCTGAGCAGTTT	1118
Db	54352	CATGGATTAATGGATATCCGATTTCTGCGTTTCAACGCGTTTGCTATCTGCCTGAGCAGTTT	54293
Qy	1119	TGAAACACAGAAATGGCTGTGAA	1140
Db	54292	TGAAACACAGAAATGGCTGTGAA	54271
RESULT 5			
BT020601			
LOCUS	BT020601 1170 bp mRNA linear PLN 05-FEB-2005		
DEFINITION	Arabidopsis thaliana At5g18680 gene, complete cds.		
ACCESSION	BT020601		
VERSION	BT020601.1 GI:58652097		
KEYWORDS	FLU CONA.		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.			
REFERENCE	1. (bases 1 to 1170)		
AUTHORS	Kim, C.J., Chen, H., Cheuk, R., Shinn, P. and Ecker, J.R.		
TITLE	Arabidopsis ORF clones		
JOURNAL	Unpublished		
REFERENCE	2. (bases 1 to 1170)		
AUTHORS	Kim, C.J., Chen, H., Cheuk, R., Shinn, P. and Ecker, J.R.		
TITLE	Direct Submission		
JOURNAL	Submitted (05-FEB-2005) Salk Institute Genomic Analysis Laboratory (SIGNAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA		

ORIGIN

Query Match	50.2%	Score 572.8;	DB 15;	Length 1170;
Best Local Similarity	71.7%;	Pred. No. 2.5e-174;		
Matches 852;	Conservative 0;	Mismatches 267;	Indels 69;	Gaps 5;
Qy	1	ATGACGTTCCGAAGCTTTACTCCAGAAATCGGCTCTAGGCACACCGCTGTAGTTTACGCC 60		
Db	1	ATGACCTTACGTAGCTTAATCTTTGAGATGGCTTCGAGACCGCATCGTGTGCTCACAAC 60		
Qy	61	-----GCGCCCTCAACCGCTAATAGTTTCAGACCCCTTTTCAGCTGCTGCG 102		
Db	61	CTTGCCGCGCGCGCAGCTGCGGATTCACATTCTGTGTTCATCGCAAGATTATCGCTGGTCA 120		
Qy	103	GAGCTCCGGAGGAGCTGCTTTAGAGAAATCTGTATTAGGTTGAGACGTGTTGACGGCGGC 162		
Db	121	GAGATTCTGAAGAGCTTCTTTAGGGAGATTCTGTATTCGTGTGAAGCGGCGGACGGTGGC 180		
Qy	163	GATTGGCGCTGCGCGCGCAAAACGTGTGGCTTGTCCGCGCTTCTGTGTAGCTGTGAGGATT 222		
Db	181	GGATGGCGCTCAGACGCGGCTGGTGGCTTGTCCGCGCTTCTGTGTAGCTGTGAGGATT 240		
Qy	223	CTCACCAAGGAGATTGTAGCTGTTCTCGAAATCTCTCTCTAAATGACTTTTCCCTATCTCC 282		
Db	241	CTTATGAACGAAACCGTCTGTTGCCCTGAGATCTCTTCTAAGTTGACTTTTCCCCCATCTCT 300		
Qy	283	CTCAAGCAGCTGTGTCCAGAGATTCTCTAGTTCAATGCTTTATAAACAAGTAATCGAAT 342		
Db	301	CTCAAGCAGCTGTGTCCAGGGATTCTCTGGTTCAATGCTTTATCAAAACGTAATCGAAT 360		
Qy	343	ACTCAATCGTATCATCTTCTCTCGGATTAACTACCTCTTTTGACGGATTAACGGGAAGTTT 402		
Db	361	ACGCAATCATCATCTCTCTCGGATTAAACCAACTCTTTAACCGATGATGGGAGTTT 420		
Qy	403	CTTCTTGCTGCTCTTAAGCTGAAACGGCGCAACTTGTGCATCTGATTAATCATCTCTTTCCGT 462		
Db	421	TTGCTTGCTGCTGTTAAGTTGAAGCACACAACTTGTACGGATTACATTATCTCTTTAGT 480		
Qy	463	TCAGACATATCTCAAGAGAGCAACGCGTATCTTGGGAGATGAGATCGAATCTCCTT 522		
Db	481	TCTGATGATATGTGCGAAGAAGCAAGCTTATGTGGCAAGTGAGATCGAACTTCCCTA 540		
Qy	523	GGAAACAAATTCAGGCTTTTGATGGTGTAGTCAG-----ACCGAGAGAGCGCAAGATG 573		
Db	541	GGAAACGAAATTCATGTCTTTTGATGGAAATCTGTGCTTCAACGGAGCGCGAAGTTG 600		
Qy	574	CAGAAGAGCGCTCTCTCTAATTTTCATCAAAGTTTTCACCTAGAGTTCTCTCAGGGAAGTTAC 633		
Db	601	AGAAAGAGCCGATCTTATATCCCGCAAGTTTCAGCAAAAGTTCTCTTTGGAAGTTAT 660		
Qy	634	CCCATCGCTCAATTTCAATAGAGTTAAACGTCTTTAGGCTCTCGGGACCGAGAAAGATG 693		
Db	661	CCTGCTGCTCATATCATATATGAGCTGAATGTCTTTAGGATCCCGGGGACCAAGAAAGATG 720		
Qy	694	CGTTGCATCATGGATACAATACCTATGAGCATCGTGGAGTCGCGAGGAGTAGTAGTTCA 753		
Db	721	CAATGTCTTTATGGACACAATACCTTCAAGGACAATGGAGCCTCAAGGAGTAGTTTCAGAA 780		
Qy	754	ACATCCATAAGCTCTTTTTCAGTTCGGTCTATCACCAAGTCTTTTAGGCTCTCACTCAAAACCA 813		
Db	781	CCATCAGAGTTTCCTTTACTCGGTACTCGGTCAACCTTATCCAGTCTCAGTCAAAACCA 840		
Qy	814	TTGCGCAGTAAATAGTGCATCATGTAGCGACTCAGGCAACAACTGGGAGATTCACCAATTG 873		
Db	841	TTACGCCAGTA-----GCTCAAGCCACCTGAAAGAAACACCACTTA 879		
Qy	874	GTGCTGAGCAACAAGCTCCACGGTGGCATGAGCAGTTACGTTGCTGGTGTCTTAAATTTTC 933		
Db	880	GTGCTGAGCAACAAGACCAACCGGTGGCAACGAGCAGCTACGCTGCTGGTGTGAATTTTC 939		
Qy	934	CATGGTCGAGTTCAGTGGGCTTCGGTTTAAAGAACTTTTCAGCTGTGTGGCAG-----TT 984		
Db	940	CATGGCCGTGTCACATGTAGGTCTAGTGAAGAACTTTTCAGCTGTGTGGCAGCAGGAGCTAGC 999		
Qy	985	AGTGACTGTGAAGCAGCGGCAGACATCTGTAGA-----GGATCATACTCTCAAGTTT 1032		

Db	1000	TGTTGGCAGTGGCCAGGGAATGTACC	GGGAGGCGAGCGAGCGAGTATATTCAGTTT	1059
Qy	1033	GGGAAAGTTGGGAGGACATGTTTACCAT	GGATATTCGATATCCGATTTCCTGCGTTTCAA	1092
Db	1060	GGGAAAGTCGGAAAGATATGTTCCAGAT	GGAATATGATACCGGATCTCAGCTTCCAG	1119
Qy	1093	GGCTTTCCTATCGCTGACGAGTTTGGAA	CCAGAAATTCGCTGTGAA	1140
Db	1120	GCTTTTGCATTTGCTTGGACAGCTTTG	AGCATAGCATGCTTGTGAA	1167
RESULT 6				
BT010768				
LOCUS				
DEFINITION				
Arabidopsis thaliana At5g18680 gene, complete cds.				
ACCESSION				
BT010768				
VERSION				
BT010768.1 GI:38564317				
KEYWORDS				
FLI CDNA..				
SOURCE				
Arabidopsis thaliana (thale cress)				
ORGANISM				
Arabidopsis thaliana				
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.				
REFERENCE				
1 (bases 1 to 1259)				
Cheuk,R., Chen,H., Kim,C.J., Shinn,P. and Ecker,J.R.				
AUTHORS				
TITLE				
Arabidopsis cDNA clones				
JOURNAL				
REFERENCE				
2 (bases 1 to 1259)				
Cheuk,R., Chen,H., Kim,C.J., Shinn,P. and Ecker,J.R.				
AUTHORS				
TITLE				
Direct Submission				
JOURNAL				
Submitted (27-NOV-2003) Salk Institute Genomic Analysis Laboratory (SIGAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA				
FEATURES				
source				
1..1259				
Location/Qualifiers				
/organism="Arabidopsis thaliana"				
/mol_type="mRNA"				
/db_xref="taxon:3702"				
/chromosome="5"				
/clone="S64136"				
/ecotype="Columbia"				
/note="This clone is in pUNI 51"				
32..1201				
Location/Qualifiers				
/notes="unknown protein"				
/codon_start=1				
/product="At5g18680"				
/protein_id="AAK23738.1"				
/db_xref="GI:38564318"				
/translation="MTLRSLILBMRSPHRVVDLAAADAADSTVSODYRWSEIPE ELUREPLRVADAGGWTSRRGVACGVRGRLMMETVVPVSSKLTPTPISLKL OPCEPLRVQCFIKRNIQTGYHLGLTNLTDDGKFLAACKLHTTCTDVIISLR SDMSRRSQAYGVKVRNPLGTFKTFVDGNLPLSTGAALKRKSRYNPAKVSAPVIG SYVAHITVELNVIGSRGPRKQCLMDTIPTSTNEPQGVASESEFPFLGTRSTLRS QXPLRSLSHLKTEPLVLSNKTPTRWHEHRCVCLNFHGVRTVASVKNFOLVAAGASC GSGTGMSPERQSERIILOFGKVGKDMFTDYGYPISAFQAFALCLSSFETRIACE"				
CDS				
ORIGIN				
Query Match				
50.2%; Score 572.8; DB 15; Length 1259;				
Best Local Similarity				
71.7%; Pred. No. 2.5e-174;				
Matches 852; Conservative				

Db 152 GAGATTCCTGAAGAGCTTCTTAGGAGGATTCGATTGCTGTTGAAGCGCGGACGGTGGC 211
Qy 163 GATTGGCCCTCGCGGCGAAACGTGTGTGCTTGTGCGCGGCTTGTGCTAGCTGGAGGATT 222
Db 212 GGATGGCCCTCAGCAGCGAGCGTGTGTGCTTGTGCGCGGCTTGTGCTGCTGGCGGCTA 271
Qy 223 CTCACCGAGGAGTGTAGCTGTTCTGTAATCTCTCTTAATCTCTTAATCTGACTTTCCTATCTCC 282
Db 272 CTTATGAACGAAACCGTGTGCTGCTGAGATCTCTTCTAAGTGAATCTTCCCTCCATCTCT 331
Qy 283 CTCACAGCAGCTGTGCTCAAGAGATCTCTAGTTCCTAATGCTTTATATAAAGCTAATCGAAT 342
Db 332 CTCACAGCAGCTGTGCTCAAGAGATCTCTGTTCAATGCTTTATCAAAGTAACTCGAAT 391
Qy 343 ACTCAATCTATCATCTCTATCTCGGATTAATCTCTCTTTCGACGATTAACGGGAGTTT 402
Db 392 ACGCAATCATATCATCTCTATCTCGGATTAACCAACTCTTTAAGCGATGATGGGAGTTT 451
Qy 403 CTTCTGTGCTCTTAAGCTGAAGCGGCAACTTGTGACTGATTAATCATCTCTTGTGCT 462
Db 452 TTGCTTGTGCTGTGTAAGTGAAGCACAACAATGTGACGATTAATCTCTTTCAGT 511
Qy 463 TCAGAGATATCTCAAGAGAGCAACGCTATCTTGGAGATGAGATCGAACTTCCTT 522
Db 512 TCTGATGATATGTCGAGAGAGCCAGCTTATGTTGGCAAGTGAATCGAATCTCTTA 571
Qy 523 GGAACAAAATTCACGGTCTTTGATGTAGTACG-----ACCGAGCAGCGAAGATG 573
Db 572 GGAACGAAATTCACGCTCTTGTGATGAAATCTGCTGCTTCAACGGGAGCGCAAGTTG 631
Qy 574 CAGAGAGCGGCTCTCTTAATTCATCAAAATTTACCTAGATGTTCTCAGGAGAGTTAC 633
Db 632 AGAAGAGCGGATCTTATAATCCGCAAAAGTTTCAGCAAAAGTTCTCTTGGAGTTAT 691
Qy 634 CCCATCGCTCACATTTTCATACGAGTTAAACGCTTAGGCTCTCGGGAGCCGAGAAATG 693
Db 692 CTTGTCGCTCATATCATATGAGCTGAATGCTTTAGGATCCGGGAGCCAGAAAGATG 751
Qy 694 CGTTGCATCATGGATCAATACCTATGAGCATCGTGGAGTCGCGAGGATGATGAGTTCA 753
Db 752 CAATGCTTTATGACACAATACCTCAAGCACAATGGAGCCTCAAGGAGTAGCTTCAGAA 811
Qy 754 ACATCATATGCTCTTTTCCAGTCGTCATACAGCTCTTTAGTCTCTACTCMAAACA 813
Db 812 CCATCAGAGTTTCCCTTACTCGGTACTCGGTCAACCTTATCCAGGCTCTCAGTCAAAACA 871
Qy 814 TTGCGCAGTAATAGTGCATCATGTAGCGACTCAGGCAACAACCTGGGAGATCCACCATG 873
Db 872 TTACGCAGTA-----GCTCAGCCACCTGAAGAAACACCATTA 910
Qy 874 GTGCTGAGCAACAAGCTCCAGGTGGCATGAGCAGTTAGCTTGTGCTGTTAAATTC 933
Db 911 GTGCTGAGCAACAAGCACCAGCGTGGCAGCAGCAGCTACGCTGCTGCTGTTGAATTC 970
Qy 934 CATGGTGCAGTACAGTGGCTTGGTTAAGAACTTTTCAGCTTTGGCAG-----TT 984
Db 971 CATGGCGGTGTCAGTAGGCTCAGTGAAGAACTTTTCAGCTGTCGCGCAGGAGTAGC 1030
Qy 985 AGTGACTGTGAAGCAGGCGCAGATCTGAGA-----GGATCATCTCCAGTTT 1032
Db 1031 TGTGGCAGTGGCAGGGAATGTACCGGAGAGCAGCAGCGGATTTATTTGCGAGTTT 1090
Qy 1033 GGGAAAGTTGGGAGGACATGTTTTACCATGATTTATGGATATCCGATTTCTGCGTTTCAA 1092
Db 1091 GGGAAAGTCGGAAAGATATGTTTCAACGATGATTTATGATACCGATCTCAGCTTTCCAG 1150
Qy 1093 GCGTTTGCATCTGCTGAGAGTTTGAACCAAGAAATTCGCTGTGNA 1140
Db 1151 GCTTTTGCCATTTGCTTGAGCAGCTTTGAGACTAGAAATCGCTGTGNA 1198

LOCUS AY046922 1143 bp mRNA linear PLN 15-APR-2004
DEFINITION Arabidopsis thaliana tubby-like protein 12 mRNA, complete cds.
ACCESSION AY046922
VERSION AY046922.1 GI:37542295
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1143)
Lai, C.P., Lee, C.L., Chen, P.H., Wu, S.H., Yang, C.C. and Shaw, J.F.
Molecular Analyses of the Arabidopsis TUBBY-Like Protein Gene
Family
Plant Physiol. 134 (4), 1586-1597 (2004)
15064372
2 (bases 1 to 1143)
Lai, C.P. and Shaw, J.F.
Direct Submission
Submitted (18-JUL-2001) Institute of Botany, Academia Sinica,
Nankang, Taipei, Taiwan 11529, Republic of China
LOCATION/Qualifiers
1. 1143
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
1. 1143
/codon_start=1
/product="tubby-like protein 12"
/protein_id="AA03978.1"
/db_xref="GI:37542296"
/translation="MRSRPHRVHDLAAARADSTSVSSODYRWEIPEELAREILIR
VEADGGWPSRRSVVACAGVCRHLLMNETVVPVEISSKLTFFISLKQKQPRSLV
QCFTKRNRTSYHLYLDTNSLDDGKFLAACKLKHTCTDYLISLRSDMSRRSQ
AYVGRVSNFLGTFTVFDGNLLPTGAALKRSYNPAKVSYPKVSQPLRSQPLRS
ELNVLGSRGPRKMOCLMPTTSTMEPOQVASEPSEFFLLGTRLSRSQPLRS
SHLKETPLVLSNKTPRHEOLRCWCLNFHGRVTVASVKNFOLVAAGASCGSGTMSPE
RQSERILLQFGVKMDMTDYGYPISAFQPAICLSSPETRIACE"
ORIGIN
Query Match 49.8%; Score 567.8; DB 15; Length 1143;
Best Local Similarity 72.8%; Pred. No. 1.1e-172;
Matches 810; Conservative 0; Mismatches 252; Indels 51; Gaps 4;
Qy 58 GCGCGCGCTCAACCGCTAATAGTTTCAGACCTTTTCAGCTGCTCGGAGCTCCGAGGAG 117
Db 49 GCTGCCGATTCACCTTCTGTGTCATCGCAAGATTATCGCTGTCAGAGATTCTCGAAG 108
Qy 118 CTGCTTAGAGAAATCTCTGATTAGGTTGAGACTGTTGACGCGCGGATTTGGCGCTCGCG 177
Db 109 CTTCTTAGGAGATTCTGATTCTGTTGAGCGGCGGACGGTGGCGATGGCGTCAACGA 168
Qy 178 CGAAACGTGGTGGCTTGTGCGCGGCTTGTCTGCTAGCTGGAGGATTTCTACCAAGAGATT 237
Db 169 CGCAGCGTGGTGGCTTGTGCGCGGCTTGTCTGCTGCTGGCGGCTACTTATGACGAAACC 228
Qy 238 GTAGCTGTTCTCGAATCTCTCTAAATTTGACTTTTCCCTATCTCCCTCAAGCAGTCTGCT 297
Db 229 GTCGTTGTCCTCGAGATCTCTTCTAAGTTGACTTTTCCCATCTCTCTCAAGCAGCTGCT 288
Qy 298 CCAAGAGATTCCTAGTTCAATGCTTTATAAAGCTAATCGAAATCTCAATCTGATCAT 357
Db 289 CCAAGGATTCATCTGTTCAATGCTTTTATCAACGTAATCGAATACGATCATATCAT 348
Qy 358 CTCATCTCGGATTAACCTACTCTTTGACGGATTAACGGGAGTTTCTTCTTGTGCTTCT 417
Db 349 CTCATCTCGGATTAACCAACTCTTTAACGGATGATGGGAAGTTTCTTGTGCTGCTGT 408
Qy 418 AGCTGAAGCGGCAACTTGTGACTGATTAATCATCTCTTGTGGTTCAGACGATATCTCA 477
Db 409 AAGTTGAAGCACACAACTTTGTACGGATTATCTTCTTTAGCTTCTGATGATATGTCG 468
Qy 478 AAGAGAAGCAACCGGTATCTTGGGAGATGAGATCGAACTTCTTCTTGGAAACAAATTCACG 537

```

Db      469  AGAAGAGCCAGCTTATGTTGCAAGTGAATCGAACTTCTTAGAACCAATTCACCT 528
Qy      538  GTCTTTGATGGTAGTCAG-----ACCGAGGACGCGAAGATGCGAAGAGCCGCTCT 588
Db      529  GTCTTTGATGGAATCTGCTGCTTCAACCGGAGCGCCAAAGTTGAGAAAGAGCCGATCT 588
Qy      589  TCTAAATTCATCAAGTTTCACCTAGAGTTCCTCAGGGAAGTTACCCCATCGCTCAATT 648
Db      589  TATAATCCCGCAAAAGTTTCAGCAAAAGTTTCTTTGGAAGTTATCTCTGCTGCATATC 648
Qy      649  TCATACGAGTTAAACGCTTTAGGCTCTCGGGGACCGAGAGAAGATGGCTTCATCATGGAT 708
Db      649  ACATATGAGCTGAATGCTTTAGGATCCCGGGACCAAGAAGATGCAATGCTTTATGGAC 708
Qy      709  ACAATACCTATGAGCATCGTGGAGTCGCGAGGAGTAGTAGCTTCAACATCAATAGCTCT 768
Db      709  ACAATACCTACAAGACCAATAGGAGCTCAAGGAGTAGCTTCAGAACCATCAGAGTTTCCC 768
Qy      769  TTTTTCAGTCGGTCATCACCACTTTTAGGTCCTCACTCAAAACCATTTGCGCAGTAATAGT 828
Db      769  TTAATCGGTACTCGGTCAACCTTTATCCAGTCTCAGTCAAAACCATTTAGCGAGTA----- 823
Qy      829  GCATCATGTAGCGACTCAGGCAACAACCTGGGAGATCCACATTTGGTGTGAGCAACAAA 888
Db      824  -----GCTCAAGCCACTGAAAGAAACCAATAGTGTGAGCAACNAG 887
Qy      889  GCTCCACGCTGGCATGAGCAGTACGTTGCTGTGCTGCTTAAATTTCCATGGTTCGAGTCACA 948
Db      868  ACACCAAGGTCGACGAGCAGCTACGCTGCTGTGCTGCTGAATTTCCATGGCCGCTGCACA 927
Qy      949  GTGGCTTCGTTAAGAACTTTAGCTTGTGGCAG-----TTAGTGACTGTGAAGCA 999
Db      928  GTAGCGTCAGTGAAGAACTTTACGTCGTGGGACGAGGAGCTAGCTGTGGCAGTGGCAGC 987
Qy      1000  GGGCAGACATCTCAGA-----GGATCATCTCCAGTTTGGAAAGTTGGGAAG 1047
Db      988  GGAATGTACCGAGAGGCGACGAGCGGATTAATTTGCGATTTGGAAAGTCGGGAAA 1047
Qy      1048  GACATGTTTACCATGGATTATGGATATCCGATTTCTGCGTTTCAAGCGTTTGTCTATCTGC 1107
Db      1048  GATATGTTCCAGATGATATGATGATACCCGATCTCAGCTTTCCAGGCTTTTGCCATTTGC 1107
Qy      1108  CTGAGCAGTTTGAACACGAAATGCTGTGAA 1140
Db      1108  TTGAGCAGCTTTGAGACTAGAATCGCTTTGAA 1140

```

RESULT 8

```

BT013601
LOCUS      1781 bp mRNA linear PLN 11-MAY-2004
DEFINITION Lycopersicon esculentum clone 132366F, mRNA sequence.
ACCESSION BT013601
VERSION    BT013601.1 GI:47105016
KEYWORDS   FLI_CDNA.
SOURCE     Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM   Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
            1 (bases 1 to 1781)
REFERENCE   Kirkness, E.P., Wang, W. and Vazeille, A.
AUTHORS    Direct Submission
TITLE      Submitted (11-MAY-2004) The Institute for Genomic Research, 9712
JOURNAL    Medical Center Drive, Rockville, MD 20850, USA
FEATURES   Location/Qualifiers
            1..1781
             /organism="Lycopersicon esculentum"
             /mol_type="mRNA"
             /db_xref="taxon:4081"
             /clone="132366F"
             /tissue_type="fruit"
             /note="TZXB151"

```

ORIGIN

```

Query Match      30.8%; Score 350.6; DB 15; Length 1781;
Best Local Similarity 59.2%; Pred. No. 6.1e-102;
Matches 678; Conservative 0; Mismatches 449; Indels 18; Gaps 4;

Qy      11  GAAGTTTACTCCAGGAAATCGCGTCTAGGCGACACCGTGTAGTTTCAGCGCGCCCTCAA 70
Db      144  GCATTTCTCGGAAGATTTGCTCCAGGTCAATAGGGTGGTTCAGGATTTACCGTGA 203
Qy      71  CGCTAATAGTTTCAGACCCCTTTTCAGCTGTGCGAGCTCCCGAGGAGCTCTTTAGAGAA 130
Db      204  CTATCGATGCGTTGAGGCAAGTTGTTGGCCCAACATGCGCGCTGAGCTATTGAGAGATG 263
Qy      131  TCTGATTAGGTTGAGACTGTTGACGGCGGCGCATTTGGCCGTCGCGCGGAAAGTGGTGG 190
Db      264  TGTGATGAGGATTTGAGGAGTCGGA---CTCTGATTTGGCGTCCAGGAAAAATTTGTTGG 320
Qy      191  CTTGTCCCGCGTTTCTGCTAGCTGAGGATTTCTCAACAGGAGATTTGTAGCTTCTCTG 250
Db      321  CGTGCCTGGTGTTCAGGAGTTTGGAGGAATCATGAAGAGATTTGTTAAACACCTG 380
Qy      251  AATTCTCTCTAAATTTGACTTTTCCCTATCTCTCCCTCAAGCAGTCTGTGTCGAAGATTTCT 310
Db      381  AGTTAGCGGAAATTTGAGCTTCCCAATCTCTCTTGAAGCAGCTGTGTCGAAGACAAAC 440
Qy      311  TAGTTCATGCTTTATAAAACGTAATCGAAATACTCAATCGTATCATCTCTATCTCGGAT 370
Db      441  TTGTTCAATGTTTTATCAGAAAGAACCGCAGCTCACAGACATATCACCTTTTCTTAATT 500
Qy      371  TAACTACCTTTTTCAGGATAACGGGAAGTTTCTCTTCTGCTGCTTCTTAAGCTGAAGCGC 430
Db      501  TAAATGAAGTTGTCTAAGATGACGTAAGTTCTCTTCTGCTCGGAAGTGTAAAGGC 560
Qy      431  CAACTTGCACATGATTAACATCATCTCTTTTGGGTTTCAGACGATATCTCAAAGAGAGCAACG 490
Db      561  CAAACATACACAGATTATCATCAATTTCTTTAAATCGCGAAGTTGCTTCCAGGGTAGCAGCA 620
Qy      491  CGTATCTTGGGAGAAATGAGATCGAACTTCTCTTGGAAACAAATTCACGGTCTTTGATGGTA 550
Db      621  CTTACATTTGGGAGCTGAGATCCAACTTCTTGGGAGCCAAAGTTTCACTGTCTATGATGCAC 680
Qy      551  GTC---AGACCGGAGCAGCGAAGATCGAGAGAGCGCTCTCTTAATTTTCATCA----- 601
Db      681  AACCATCTAATGCAGGAGCTTAAAGTCAGTAAATGTCTGCTCTACTAGTTTGTGGGATGA 740
Qy      602  ---AAGTTTTCACCTAGAGTTTCTTCAGGGAAGTTTACCCCATCGCTCACTTTCATACAGAT 658
Db      741  AACAGTCTCTCAAGAAATACCACTGGCAACTATCTCTGTAGCTCACATCTCGTATGAT 800
Qy      659  TAAACGCTTTAGGCTCTCGGGACCGAGAAAGATGCGTTGATCATGATGATGATCAATACCTA 718
Db      801  TGAATGCTCTGGGGCTAGGGGGCCCGAGGAGATGCTGTGTATCATGAGCAATTCAG 860
Qy      719  TGAGCATCTGGAGTCGCGAGGAGTAGTAGCTTC---AACATCCATAAGCTCTTTTCCA 775
Db      861  CTTCTTTCTAATGAAACCTGGGGGTTTAGCCCTTACGCAACAGAAATTTGTTCCAGTGAATG 920
Qy      776  GTCGGTCAATCACCAGCTTTTAGTGTCTCACTCAAAAACCATTTGCGCAGTAATAGTGCATCAT 835
Db      921  TAGATTTCTTCCCTCCCTCCACTTTTGTAGATCAAAATCAAACTCGGTGGATTAATTTCT 980
Qy      836  GTAGCGACTCAGGCAACAACTCTGGGAGATCCACCATTTGGTGTGAGCAACAAAGCTCCAC 895
Db      981  CATCCGGGCTTTTACCAGTCCAAAAGATGAAGTTCTTAATTTGAAAGAACAAAGCTCCTA 1040
Qy      896  GGTGGCATGAGCAGTTAGCTGTGCTTGTGCTTAAATTTCAATGTCAGTGCAGTGGCTT 955
Db      1041  GGTGGCATGAACAACTCCCAATGCTGTGCTTAACTTTCAATGTCGCGGGTGAAGTGGCTT 1100
Qy      956  CGGTTAAGAACTTTTACGCTTGTGGCAGTTAGTCACTGTGTAAGCAGGCGACACATCTGAGA 1015
Db      1101  CTGTAAAAAATCTTTTTCAGTAGTTGCTTCTCTTGGGATGGAGCTGGACAGGAAATGAGA 1160

```


QY 1016 GGATCATACTCCAGTTTGGAAAGTTGGAGGACATGTTTACCATGGATTATGATATC 1075
Db 1161 ATGTTATCTCCAAATTTGGAAAGTGGAGAGATTTGTTCCATGGACTATCAGTATC 1220
QY 1076 CGATTTCCTCGTTTCAAGCGTTTGTCTATCTGCTGAGCAGATTTTGAACCCAGAAATTCCT 1135
Db 1221 CTATCTCTGCATTTCAAGCATTTGCTATCTGCTCAGTAGCTTTGACACCAAAATCGCCT 1280
QY 1136 GTGAA 1140
Db 1281 GCGAA 1285
RESULT 9
CQ804888
LOCUS CQ804888 1221 bp DNA linear PAT 10-MAY-2004
DEFINITION Sequence 1299 from Patent WO2004035798.
ACCESSION CQ804888
VERSION CQ804888.1 GI:47111078
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1
AUTHORS Inze, D., de Veylder, L. and Vlieghe, K.
TITLE Identification of novel e2f target genes and use thereof
JOURNAL Patent: WO 2004035798-A 1299 29-APR-2004;
CropDesign N.V. (BE)
FEATURES
source 1..1221
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"
ORIGIN
Query Match 30.1%; Score 342.6; DB 6; Length 1221;
Best Local Similarity 59.8%; Pred. No. 2.4e-99;
Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;
QY 32 GGTCTAGGGCCACACCGTGTAGTTTCACGCCGCCCTCAACCGCTAATAGTTCAGACCCCTT 91
Db 89 GATCCAGGTTCTCAACGTGTGTTTCAGAGATCTTCTGTTCTGTGTTGATGCTTTCAACGACA 148
QY 92 TCAGCTGGTTCGGAGCTCCCGAGGAGCTGCTTAGAGAAATCCCTGATTAGGGTTGAGACTG 151
Db 149 GCTGCTGGGCTAGTATGCTCCGGAGCTCTCGAGAGATGTTCTTATGAGGATTGAGCAAT 208
QY 152 TTGACGGCGCGGATTTGGCGTCCGGCGGAAACGTGGTGGCTTTGCGCGGTTTGTTCGTA 211
Db 209 CCGAAGACAC---TTGGCGCTCTAGGAAATAATGTTTCTTGGCTGTGTCTGCAGGA 265
QY 212 GCTGGAGGATTCACACAGGAGATGTAGCTGCTCTGTAATCTCTCTAAATTCACATT 271
Db 266 ACTGGCGAGAAATCGTCAAGAGATCGTCAGATCTCTGAGCTTCTGAGCTTTCTAGCAAACTCAC 325
QY 272 TCCTATCTCCCTCAAGCAGCTGTGGTCCAGAGATCTCTAGTTCATGCTTTTATAAAAC 331
Db 326 TTCCTATCTCCCTCAACAGCCGGTCTTAGAGATCACTTGTTCATGCTATATTATGA 385
QY 332 GTAATCGAAATATCAATGTAATCTATCTCGGATTAA---CTACCTCTTTGACGG 388
Db 386 GAAACCGCAGCAATCAAACTACTATCTATACCTCGGGTTAAACCAAGCAGCTTCAAATG 445
QY 389 ATAAACGGAGTTTCTTCTGCTGCTTCTTAAGCTGAAGCGCAACTTGCACCTGATTACA 448
Db 446 ATGATGGAAGTTCTTCTTGTCTGCCAAGAGGGTTTCGGAGGCCAACTTGCACCTGACTACA 505
QY 449 TCATCTCTTTGCGCTTCAGACGATATCTCAAGAGAGAACGCGGTATCTTTGGAGAGATGA 508
Db 506 TCATCTCTTAACCTGCGATGATGCTCTCTCGAGGAAGCAATACCTATATCGGAAAGCTTA 565

QY 509 GATCGAACTTCCTTGGAAACAAATTTCAAGTCTTTTGATGGTAGTCAGACCGGAGCAG--- 565
Db 566 GATCTAACTTTCTGGGACCAAGTTCACTGTCTATGACGCTCAGCCGCAAGTCTCTGGAA 625
QY 566 -----CGAAGATGAGAGAGAGCCGCTCTTTCTAAATTTTCATCAAAAGTTTCACTTAGAG 616
Db 626 CTCAGGTTTACAGAACCCGTTTCAAGCAGACTTCTCAGTTTGAACAAGTGAGCCCGAGAA 685
QY 617 TTCCTCAGGAAGTTACCCCATCGCTCAATTTTCATACGAGTTAAACGTTCTTAGGCTCTC 676
Db 686 TTCCATCTGGCAACTATCTCTGTAGCACATATCTATAGAGCTTAAACGCTTTGGGTTCCA 745
QY 677 GGGGACCGAGAGAAATCGTTTCATCATGGATACAAATACCTATGAGCATCGTGGAGTCGC 736
Db 746 GAGGACCGAGGAGGATGAGTGTGTCATGGATGCCATCCCTGCATCAGCTGTAGAACCTG 805
QY 737 GAGGAGTAGTAGTCTTCAACATCCATAAGCTCTTTTTCAGTC-----GGTCATCACCA 790
Db 806 GAGGAACAGCTCCAACCTCAGACGGAACTTGTCCATAGCAATCTTGATAGTTTCCCTCAT 865
QY 791 TCTTTAGTCTCAGTCAAAACCAATTCGCGAGTAAATAGTCATCATGTAGCAGTACGAGTAC 850
Db 866 TCTCCTTCTTCAAGTTCGAAATCAATTCGTCGAGAGATCTCCCTTCTGTCATCATCTG 925
QY 851 ACAACCTGGAGATCCACCATTTGCTGAGCAACAAGCTCCAGCTCGGTCGATGAGCACT 910
Db 926 CTGCTCAGAGAGGAGATGCTGTTGCTGAAACCAAGGCCAGATGGCAGCAACAGC 985
QY 911 TACGTTGCTGCTGCTTAAATTTCCATGTCGAGTCACAGTGGCTTCGGTTAAGAACTTTC 970
Db 986 TCCAGTGTGCTGCTCAACTTCAATGGAGAGTCACAGTTGCTTCCGTCGTCGTCGTCGTC 1045
QY 971 AGCTTGTGCGAGTGTAGTACTGTGAA---GAGGGCAGACATCTGAGAGATCATACTCC 1027
Db 1046 AGCTGCTAGTCTCTCTCTGAGAAATGGACCTGCGAGGACCTGAGCAGCAACGATTCCTC 1105
QY 1028 AGTTTGGGAAAGTTGGAGGAGATGTTTACCATGATGATGATATCGATTTCGCTTCTGCT 1087
Db 1106 AGTTTGGAAAGTCGAAAGATGTTTCAATGATGATGATGATGATGATGATGATGATGATG 1165
QY 1088 TTCAAGCGCTTCTATCTGCTGAGCAGTTTGAACCAAGAAATTCGCTGTGAA 1140
Db 1166 TCCAGGCTTCACTTTCCTTCCCTCAGCAGTTTCGACACCAAGATGATGATGAA 1218
RESULT 10
AY045774
LOCUS Arabidopsis thaliana 1221 bp mRNA linear PLN 15-APR-2004
DEFINITION Arabidopsis thaliana tubby-like protein 3 mRNA, complete cds.
ACCESSION AY045774
VERSION AY045774.1 GI:27372513
KEYWORDS Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 1221)
AUTHORS Lai, C.P., Lee, C.L., Chen, P.H., Wu, S.H., Yang, C.C. and Shaw, J.F.
TITLE Molecular Analyses of the Arabidopsis TUBBY-Like Protein Gene Family
JOURNAL Plant Physiol. 134 (4), 1586-1597 (2004)
PUBMED 15064372
REFERENCE 2 (bases 1 to 1221)
AUTHORS Lai, C.P. and Shaw, J.F.
TITLE Cloning and characterization of cDNAs for tubby-like protein 3
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 1221)
AUTHORS Lai, C.P. and Shaw, J.F.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-2001) Institute of Botany, Academia Sinica, Nankang, Taipei, Taiwan 11529, Republic of China

FEATURES	source	Location/Qualifiers
		1..1221
		/organism="Arabidopsis thaliana"
		/mol_type="mRNA"
		/db_xref="taxon:3702"
CDS		1..1221
		/notes="putative Tub family protein"
		/codon_start=1
		/product="tubby-like protein 3"
		/protein_id="AAK98802.1"
		/db_xref="GI:27372514"
		/translation="MSFSKLIQDMGELGSLSRKGFVDFPGYGRSRORVVQDTSVPV DAFQSCASMPPELLRDIQESDQETWPRKNVWSCAGVCRNWRIVKEIVRVE LSKUTFTYISLQCDVSGSLVQCYIMRNSQTYLYLGLNQAASNDKFLVAKRF RRPTCTDIYISLNCDDVSGSNYYIGKLSNPLFTKFTVYDQAQPTNPGQVTRRSR ELVSHKSPRIISGNYPVAVHISYELNLSGPRMOCMDAIPASAVEPGCTAPTQT LILVSHNLSDFSPFSRFSIRAESLPSGPSAAQKEGLLVLNKAPRWHEOLCWCIL NENGRVTVASVKNFQLVAEPNGPAGPHENVILQFGVKGVDMYQYPIISAPQAF TICLSDFDKIACE"
ORIGIN		
	Query Match	30.1%; Score 342.6; DB 15; Length 1221;
	Best Local Similarity	59.8%; Pred. No. 2.4e-99;
	Matches	677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;
Qy	32	GGTCTAGGCACACCGTGTAGTTTCAGCGCGCGCTCAACCGCTAATAGTTTCAGACCCGTT 91
Db	89	GATCCAGGTCCTCAACGTGTTGTTTCAGGATACCTCTGTTCTGTTGATGCTTTCAAGCAGA 148
Qy	92	TGAGCTGTTCGGAGCTCCCGGAGGAGCTGCTTAGAGAAATCCTGATTAGGTTGAGACTG 151
Db	149	GCTGCTGGGCTAGTAGTGCCTCGGAGCTCCTCAGAGATGTTCTTATGAGGATTGAGCAAT 208
Qy	152	TTGACGCGCGGAGTTGCGCGTGCAGCGGAAACGTGTGCTTGTGCGCGGCTTTGTGCTGA 211
Db	209	CCGAAGACAC---TTGGCCGCTTAGGAAANAATGTTGTTTCTTGGCCGTGTCTGCGAGGA 265
Qy	212	GCTGGAGGATTCCTCAACCAAGGAGATTGAGCTGTGTTCTCGAATTCCTCTCTAAATTGACTT 271
Db	266	ACTGGCGAGAAATCGTCAAGAGATGCTCAGAGATTCTCGAGCTTTCTAGCAAACTCACTT 325
Qy	272	TCCTATATCCCTCAAGCAGTCTGTGTCGAAGATGTTCTAGTTCGATGCTTTTATAAAC 331
Db	326	TTCTCTATCTCCCTCAACACGCGGGTCCCTAGAGATGTCATCTGTTTCAATGCTATATTGA 385
Qy	332	GTAATCGAATATCTCAATCGTATCTCTATCTCGGATTAAC---CTACCTCTTTGACGG 388
Db	386	GAAACCGCAGCAATCAAACTTACTATATACCTCGGTTAAACCAACGAGCTTCAATG 445
Qy	389	ATAACGGGAAGTTCTCTTCTGCTGTTCTTAAGCTGAACGCGCACTTTGCTACTGATACA 448
Db	446	ATGATGGAAGTTCTCTTCTGCTGCTGCAAGAGGTTTCGAGGCCAACTTGCTACTGACTACA 505
Qy	449	TCATCTCTTTGGGTCAGACGATATCTCAAGAGAACCAACGCGTATCTTGGGAGAATGA 508
Db	506	TCATCTCTTAACTCCGATGATGTTCTCTCGAGGAAGCAATACCTATATCGGAAAGCTTA 565
Qy	509	GATCGAACTCTTTGGAACAAAATTCACGGTCTTTGATGGTAGTCAGACCCGAGCAG--- 565
Db	566	GATCTAACTTTCTGGGGACCAAGTTTCACTGTCTATGACGCTCAGCCGACGAATCCTGGAA 625
Qy	566	-----CGAAGATGCAAGAGCGCGCTCTTCTTAATTTCATCAAGTTTTCACCTAGAG 616
Db	626	CTCAGGTTACGAGACCCGTTTCAGCAGACCTTCTCAGTTTGAACAGTAGGACCGCAGAA 685
Qy	617	TTCCCTCAGGGAAGTTACCCCATCGCTCACAATTTCTATACAGGTTTAAACGCTCTTAGGCTCTC 676
Db	686	TTCCATCTGGCAACTATCTGTAGCACATATCTCATATGAGCTTAAACGCTTTGGGTTCCA 745
Qy	677	GGGGACCGAGAGATGCGGTTTGATCATGGAATACAAATCCTATGAGCATCGATCGTAGTGGC 736
Db	746	GAGGACCGGAGGATGTCAGTGTGTGTCATGGATGCCATCCCTCGCATCAGCTGTAGAACCTG 805

Banh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.	
Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC) contributed equally to this work as PIs.	
Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.	
FEATURES	Location/Qualifiers
source	1. .1252
	/organism="Arabidopsis thaliana"
	/mol_type="mRNA"
	/db_xref="taxon:3702"
	/chromosomes="2"
	/clone="U19776"
	/ecotype="Columbia"
	/note="This clone is in pUNI 51."
gene	1. .1252
	/gene="At2g47900"
CDS	1. .1221
	/gene="At2g47900"
	/codon_start=1
	/evidence=experimental
	/product="putative tubby protein"
	/protein_id="AA020254.1"
	/db_xref="GI:20465543"
	/translation="MSFKSLIQMRGELGISRKGFDPVFGYGRSRSQRVVQDTSVPVDAFKQSWAMPPELLRDVLMRIEQLSEDTWPSKKNVSCAGVCRNWRREIVKSVRPELSKLTFFLSLQPGGRGLVQCYINRNSNOTLYLYLGLNQAANDGKFLLAARFRRPTDYIISLNCDDVSRGSNTYIIGKLSNLFYDQTPNPGTQVTRSSRLSLKQVSRIPDSGNYPVAHISVELNLGSRGPRMQCVMDAIPASAVEPGGTAPTQT ELVSNLSDSPSPSPRSKISRAESLPSGSSAAQKEGLLVLNKAPRWHEOLQWCLNFGNRTVASVKNFQVLAAPENGPAGPEHENVILQPGKVGKDFVMDYQYPISAQAF TICLSFDFTKIACE"
3' UTR	1222. .1252
	/gene="At2g47900"
ORIGIN	
Query Match	30.1%; Score 342.6; DB 15; Length 1252;
Best Local Similarity	59.8%; Pred. No. 2.4e-99;
Matches	677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;
QY	32 GGCTAGGCCACCGCTGTAGTTTACGCGCGCGCCCTCAACCGCTAATAGTTTCAGACCCCTT 91
DB	89 GATCCAGGCTCAACGCTGTGTTTCAGGATACTTCTGTTCTGTTGATGCTTTCAGCAGA 148
QY	92 TCAGCTGGTCGGAGCTCCCGAGAGAGCTGCTTAGAGAAATCCTTGATTAGGTTGAGACTG 151
DB	149 GCTGCTGGGCTAGTATGCTCCGAGCTCTCGAGATGTTCTTATGAGGATTTGAGCAAT 208
QY	152 TTGACGGCGCGATTGGCCGTCGCGGGRACCTGGTGGCTTGTGCGGGGTTTGCCTA 211
DB	209 CCGAAGACAC---TTGGCGGCTTAGGAAAATGTTGTTCTTGGCTGGTGTCTGCAGGA 265
QY	212 GCTGGAGGATTTCCACCAAGGAGATTGTAGCTGTCTGAATTTCTCTCTAAATTGACTT 271
DB	266 ACTGGGAGAAATCGTCAAGAGATCGTCAGAGTTCTGAGCTTTCTAGCAAACTCACT 325
QY	272 TCCTATCTCCCTCAAGCAGTCTGGTCCCAAGAGATTCTCTAGTTCAATGCTTTATAAAC 331
DB	326 TTCTATCTCCCTCAACAGCGGGCTTAGAGATCACTTGTTCATATGCTATATTATGA 385
QY	332 GTAATCGAATATCAATCGTATCACTCTATCTCGGATTAA---CTACCTCTTTACCG 388
DB	386 GAAACCGCAGCAATCAAACTCTATCTATACCTCGGGTTAAACCAAGCAGCTTCAAATG 445
QY	389 ATAAACGGAGATTCTTCTTGTGCTTCTAAGCTGAAGCGCGCAACTTGCACCTGATTACA 448

DB	446 ATGATGGAAGATTCTTCTTGTGCTGCAAGAGGTTTCGGAGGCCAACTTTCGCACTGACTACA 505
QY	449 TCATCTCTTTGGCTTCAGACGATATCTCAAAGAGAGCAACGCGTATCTTTGGGAGATGA 508
DB	506 TCATCTCTTAACTGCGCATGATGCTCTCGAGGAGCAATACCTATATTCGGAAGACTTA 565
QY	509 GATCGAACTTCTCTTGGAAACAAAATTCACGGCTTTTGTATGGTAGTCAGACCGGAGCAG--- 565
DB	566 GATCTAACTTTCTGGGACCAAGTTCACTGTCTATGACGCTCAGCCGAGAACTCTGGAA 625
QY	566 -----CGAGATGACAGAGAGCGCTCTCTTCTAAATTTCTATCAAGTTTCACTAGAG 616
DB	626 CTCAGGTTTACCAAGAACCCGTTTCAAGCAGACTTCTCAGTTTGAACAAGAGTAGCCCGAGAA 685
QY	617 TTCCTCAGGAGAGTTTACCCCATCGCTCAATTTTCATACGAGTTTAAACGCTTTAGGCTCTC 676
DB	686 TTCCATCTGGCAACTATCTCTGAGCACAATCTCATATGAGCTTAAACGCTTTGGGTTCCA 745
QY	677 GGGGACCGAGAGAAATGCGTTGTCATCATGGATFACAAATACCTATGAGCATCGTGGAGTCGC 736
DB	746 GAGGACCGGAGGATGAGTGTGTCATGATGCCATCCCTGCATCAGCTGTAGAACCCTG 805
QY	737 GAGGAGTAGTAGTTCAACATCATAAGCTCTTTTTCAGATC-----GGTCATCACAG 790
DB	806 GAGGAACAGCTCCAACTCAGACGGAACCTTGTCCATAGCAATCTTGATAGTTTCCCTCAT 865
QY	791 TCTTTAGTCTCACTCAAAACCATTTGCGCAGTAATAGTGCATCATGTAGCGACTCAGGCA 850
DB	866 TCTCTTCTTCAGGTCGAAATCAATTCGTGACAGAGTCTCCCTCTTGTGTCATCATCTG 925
QY	851 ACAACTCGGAGATCCACCATTTGGTGTGAGCAACAAAGCTCCACGGTGGCATGAGCACT 910
DB	926 CTGCTCAGAGGAGAGACTGTTGCTCTGAAAAACAAGCGCCAGATGGCAGACAGC 985
QY	911 TACGTTGCTGGTCTTAAATTTCCATGTCAGTCAAGTGGCTTCGGTTTGAAGAACTTTC 970
DB	986 TCCAGTGTGGTGCCTCAACTTCAATGGAGAGTCAAGTTGCTTCCGTCACAAAACTTTC 1045
QY	971 AGCTTGTGGCAGTTAGTACTGTGAA---GCAGGCGACACATCTGAGAGGATCATACTCC 1027
DB	1046 AGCTGTGTAGTCTCTCTGAGAAATGGAACCTGCGAGGACCTGAGCAGCAAAACGTTCTCC 1105
QY	1028 AGTTTGGGAAAGTTGGGAGGACATGTTTACCATGGATTATGGATATCCGATTTCTGCGT 1087
DB	1106 AGTTTGGGAAAGTCGAAAAAGATGTTTCAATGGATTATCAGTACCCTATCTGCTCT 1165
QY	1088 TTCAAGCGTTTGCTATGCTGCTGAGCAGTTTGTGAAACCGAAGATTGCTCTGAA 1140
DB	1166 TCCAGGCTTTCACCATTTGCTTCCAGAGTTTCGACACCAAGATAGCATGTGAA 1218
RESULT 12	
AY074273	1943 bp mRNA linear PLN 18-SEP-2002
LOCUS	Arabidopsis thaliana putative tubby protein (At2g47900) mRNA,
DEFINITION	complete cds.
ACCESSION	AY074273
VERSION	AY074273.1 GI:18377639
KEYWORDS	FLI CDNA.
SOURCE	Arabidopsis thaliana (thale cress)
ORGANISM	Arabidopsis thaliana
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
	rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE	1 (bases 1 to 1943)
AUTHORS	Yamada, K., Liu, S.X., Sakano, H., Pham, P.K., Banh, J., Chung, M.K., Goldsmith, A.D., Lee, J.M., Quach, H.L., Toriumi, M., Yu, G., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayaishizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.
TITLE	Arabidopsis Full Length cDNA Clones

JOURNAL REFERENCE AUTHORS	Unpublished 2 (bases 1 to 1943) Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,B., Dale,J.M., Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,J.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A. Direct Submission Submitted (11-JAN-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.	Db
TITLE JOURNAL COMMENT	The Salk, Stanford, PGENC (SSP) Consortium members carried out the sequencing and annotation of the RAPL cDNAs: Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,B., Dale,J.M., Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,J.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Chen,H., Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Davis,R.W., Ecker,J.R. and Theologis,A. Yamada,K. (SSP/PGENC) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGENC) contributed equally to this work as PIs. Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank. Location/Qualifiers 1. .1943 /organism="Arabidopsis thaliana" /mol_type="mRNA" /db_xref="taxon:3702" /chromosome="2" /clone="RAPL09-89-B17 (R19776)" /ecotype="Columbia" /note="This clone is in a modified pBluescript vector (pLC-1) as a BamHI/XhoI insert." 1. .1943 /gene="At2g47900" 1. .454 /gene="At2g47900" 455. .1675 /gene="At2g47900" /codon_start=1 /evidence=experimental /product="putative tubby protein" /protein_id="AA166970.1" /db_xref="GI:18377640" /translation="MSFKSLIQDMRGELGISRKGFVDFVGYGRSRQRVQDTSVPVDAFKQSCAMPELLRDVIMRIEISEDWPSKONVSCVGRNWRREIVKGIVRYPE LSSKLTFFISLQPGPRGSLVQCIYMRNSQTYLYILGNQASNDKGFLLAAKRF RPTCTDYIISLNCDDVSRSGTYIGKLSNFGTKFTVYDAQPTNPQVTRTRSSR LLSLKQVSPRIPSGNPVAHISVELNVLGRGPRMQCVMDALPASAVEPGTAPTOT ELVHNLDSFPSPFSRKSIRAESLPSGSPSSAAQKEGLLVLNKAPRWHEQLQWCL NFNGRVTVASVKNFQLVAAPENGPAGPEHNVILQFGRVKDVFNDIYPIPSAQFAT TILCSFDFTKIACE" 1676. .1943 /gene="At2g47900"	Qy
3' UTR	Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5; Qy 32 GGTCTAGGCCACCGTGTAGTTTCAGCCGCGCCTCAACCGCTAATAGTTTACAGCCCTT 91	Qy
ORIGIN	543 GATCCAGGTCTCAACGCTGTGTTTTCAGGATACCTTCTGTTCTCTGATGCTTTCAAGCAGA 602 92 TCAGCTGTGGGAGCTCCCGAGGAGCTGCTGTAGAGAAATCTCGATTAGGGTTGAGACTG 151 603 GCTGCTGGGCTAGTATGATGCTCCGGAGCTCTCGAGAGATGTTCTTATGAGGATGAGCAAT 662 152 TTGACGGCGGCGATTGGCGCTGCGCGGCAAAACGCTGGTGGCTTGTGCGGCGGTTTGTGCTA 211 663 CCGAAGACAC---TTGGCCGCTTAGGAAAATGTTGTTTCTTGGCTGGTGTCTGCGAGA 719 212 GCTGGAGGATCTCACCAAGGAGATGTAGCTGTTTCTGAAATCTCTCTAAATGACTTT 271 720 ACTGGCGAGAAATCGTCAAGAGATCGTCAGAGTTCCTGAGCTTCTAGCAAACTCACTT 779 272 TCCCTATCTCCCTCAAGAGCTGTGGTCCCAAGAGATTTCTCTAGTTCAATGCTTTTATAAAC 331 780 TTCCTATCTCCCTCAAAACAGCCGGTCTCTAGAGGATCACTTGTTCATGCTATATTATGA 839 332 GTAATCGAAATACTCAATCGTATCATCTCTATCTCGGATTAA---CTACCTCTTTGACGG 388 840 GAAACCGCAGCAATCAAACTTACTATCTATCTCTCGGGTTTAAACCAAGCAGCTTCAATG 899 389 ATAACGGGAAAGTTTCTTCTTGTGCTTCTTAAGCTGAAGCGCGCAACTTGCACCTGATTACA 448 900 ATGATGGAAGTTTCTTCTTGTGCTGCGCAAGAGGTTTCGGAGGCCAACTTGCACCTGACTACA 959 449 TCATCTCTTTGGTTTCTGAGAGATATCTCAAGAGAGCAACGGTATCTTTGGGAGATGA 508 960 TCATCTCTTAAACTGCGATGATGCTCTCGAGGAAGCAATACCTATATCGGAAAGCTTA 1019 509 GATCGAACTTCTTGGCAAAAATTCACGGTCTTTCATGCTGTAGTGTAGTCAGACCGGAGCAG-- 565 1020 GATCTAACCTTCTGGGACCAAGTTTCACTGTCTATGACGCTCAGCGCAGCAATCTCTGGAA 1079 566 -----CGAAGATGCAAGAAGAGCGCTCTTTCTAATTTTCATCAAGTTTTCACCTAGAG 616 1080 CTCAGGTTACAGAACCCGTTCAAGCAGACTTCTCAGTTTGAAGAGTGAGGCCGAGAA 1139 617 TTCCTCAGGGAAGTTTACCCCATCGCTCACAATTTTCATACGAGTTAAACGCTTTAGGCTCTC 676 1140 TTCCATCTCGCAACTATCTCTGATGACATATCTCATATGAGCTTAACTGCTTTGGGTTCCA 1199 677 GGGGACCGAGAGAAATCGTTTCATCATGATACATACATATGAGCATCGTGGAGTCCG 736 1200 GAGGACCGAGGAGGATGAGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1259 737 GAGGAGTAGTAGCTTCAACATCCATAAGCTCTTTTTCAGTCT-----GCTCATCAACCAG 790 1260 GAGGAACAGCTCCAACCTCAGACGGAATCTGTCCATAGCAATCTTGATAGTTTCCCTCAT 1319 791 TCTTTAGGTCTCACTCAAAACCAATTCGCGAGTAATAGTGCATCATGTAGCGACTCAGGCA 850 1320 TCTCTCTTCTCAGGTCGAATCAATTCGTGTCAGAGAGTCTCCCTTCTGGTCCATCATCTG 1379 851 ACACCTGGAGATCCACCAATTTGGTCTGTCAGCAACAAAGCTCCACGGTGGCATGAGCAGT 910 1380 CTGCTCAGAAAGAGAGACTGCTGTGCTGAAACAAAGCGCCAGATGGCAGCAACAGC 1439 911 TACGTTGTGCTGCTTAAATTTTCCATGCTGTCAGTCACTGCTGCTTCCGTTTAAAGAACTTTC 970 1440 TCCAGTGTGCTGCTCACTTCAATGGAGAGTCACTGCTGCTTCCGTCGTCGTCGTCGTCGTC 1499 971 AGCTTGTGTCAGTGTAGTCTGTGAA---GCAGGGCAGACATCTGAGAGGATCATACTCC 1027 1500 AGCTGTAGTCTCTCTGAGAAATGGACCTGCAGGACCTGAGCAGCAAAACGCTGATTTCTCC 1559 1028 AGTTTGGGAAAGTTGGGAGGACATGTTTACCATGATTTTGGATATCCGATTTCTCGCT 1087 1560 AGTTTGGGAAAGTGGGAAAGATGTTTCCCAATGGATTTATCAGTACCTTCTCTGCTCT 1619 1088 TTCAAGCGTTTCTCTCTCTGCTGAGCAGTTTGGAAACCAAGAAATTCGCTCTGTGAA 1140 1620 TCCAGGCTTTCACCTTTGCTCTCAGCAGTTTTCGACCAAGATAGCATGTGAA 1672	Db

RESULT 13

AK071159

LOCUS

DEFINITION

AK071159 1687 bp mRNA linear PLN 24-JUL-2003

Oryza sativa (japonica cultivar-group) cDNA clone:J023078015, full

insert sequence.

ACCESSION

AK071159

VERSION

FLI CDNA; CAP trapper.

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

The Rice Full-Length cDNA Consortium, National Institute of
Agrobiological Sciences Rice Full-length cDNA Project Team:
Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,
Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,
Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,
Ohta, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,
Kurotsaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, O., Lu, M.,
Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J.,
Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,
Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN,
Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,
Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,
Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y.,
Saito, R., Sasaki, D., Sato, K., Shibata, K., Shingawa, A., Shiraki, T.,
Yoshino, M., and Hayashizaki, Y.

Collection, mapping, and annotation of over 28,000 cDNA clones from

Japonica rice

Science 301 (5631), 376-379 (2003)

12869764

2 (bases 1 to 1687)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,
Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,
Hayaehida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T.,
Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K.,
Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,
Kawagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,
Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,
Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,
Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,
Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,
Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K.,
Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H.,
Osato, N., Ota, Y., Ootomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K.,
Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,
Shingawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,
Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,
Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A.,
Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W.,
Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and
Yoshimura, A.

Direct Submission

Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of

Agrobiological Sciences, Department of Molecular Genetics, Head of

Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki

305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp,

Tel: 81-29-838-7007, Fax: 81-29-838-7007)

This clone is one of the 28K full-length cDNA clones from japonica

rice.

URL : <http://cdna01.dna.affrc.go.jp/cDNA/>

NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K.,

Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,

Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T.,

Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and

Yamamoto, M.

TITLE

JOURNAL

COMMENT

FAIS Genome Sequencing & Analysis Group: Ootomo, Y., Iida, Y.,
Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,
Kodama, T., Kurotsaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J.,
Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S.,
Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,
Yoshimura, A., Matsubara, K. and Murakami, K.
and Genome Science Laboratory in Riken Genomic Sciences Center
Genome Exploration Research Group in Riken Genomic Sciences Center
Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,
Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,
Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y.,
Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,
Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Ootomo, N.,
Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,
Sasaki, D., Sato, K., Shibata, K., Shingawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,
Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,
Yasunishi, A. and Hayashizaki, Y.

FEATURES

source

1..1687

/organism="Oryza sativa (japonica cultivar-group)"

/mol_type="mRNA"

/cultivar="Nipponbare"

/db_xref="taxon:39947"

/clone="J023078015"

ORIGIN

Query Match

Best Local Similarity

Matches

Conservative

Mismatches

Indels

Gaps

3;

18;

152

409

212

466

272

526

332

586

392

646

452

706

512

766

565

826

620

886

680

946

Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,
Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N.,
Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,
Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, P.,
Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,
Yasunishi, A. and Hayashizaki, Y.

FEATURES

source
1. .2102
Location/Qualifiers
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J033091L01"

ORIGIN

Query Match 23.1%; Score 262.8; DB 15; Length 2102;
Best Local Similarity 58.3%; Pred. NO. 2.5e-73;
Matches 506; Conservative 0; Mismatches 347; Indels 15; Gaps 2;
QY 288 GCAGTCTGCTCAAGAGATTCTCTAGTTCAGTCTTATAAAGCTTAATCGAAATACTCA 347
DB 24 GCAGCTGCCCCAAGAGCGCACTCTTAATGTTTCATAGAGAAACGGACTACTCA 83
QY 348 ATCGTATCATCTCTATCTCGATTAACTACCTCTTTGACGATPAACGGGAAGTTCTTCT 407
DB 84 GACATATTATCTGTATATTGGGCTCACAGAAGCACTTGGCGATGATGGGAAGTTCTTACT 143
QY 408 TGCTGCTTCTAAGCTGAAGCGGCACTTGGCACTGATTACATCATCTCTTTGGTTCAGA 467
DB 144 TCGTGCACCAATGACGAAGCCCATGCAAGACTACCTGATTTCCTTTGACATGAG 203
QY 468 CGATATCTCAAAGAGAGCAACGCGTATCTTGGGAGATGAGATGAACTTCTCTTGGAAC 527
DB 204 TGTATGTCAAGGGGAGGCAACACTATATTGGCAAGCTAAGGTCAAACTTTCTCGAAC 263
QY 528 AAAATTCACGGTCTTTGATG-----GTAGTCAGACGGGAGCAGCGAAGATGCA 575
DB 264 AAAGTTTACGGTCTATGATGCCATCCACATATGATGAGCTGTTGTCTCAAAAAGTCG 323
QY 576 GAAGAGCCCTCTTCTAATTTTCATCAAGTTTCACTAGATTTCTTCAGGGAAGTTACCC 635
DB 324 GTCTGCACGTGTGGTGTGTAACCAAGTCTCCCGAGAGTTCTTGCCTGGGAATTTACCC 383
QY 636 CATCGCTCACATTTTCATACGAGTTAAAGCTCTTAGGCTCTCGGGACCGAGGAAGATCG 695
DB 384 AGTTTCACATATTTCTACGAATTTGAATGTTTGGGTGCAAGAGTCCCAAGAGGATGAA 443
QY 696 TTGCAATCATGGATACAAATACCTATAGCATCTGTGGAGTCGCGAGGAGTAGTAGCTTCAAC 755
DB 444 CTGCATTATGGATTCCATCCCAACATCAGCTGTTCAGAGGGTGGGAAAGCTCCGACACA 503
QY 756 ATCCATAAGCTCTTTTTCAGTGGTTCATCAACAGTCTTTAGGTCTCATCTCAAAACCAT 815
DB 504 GACTGAATTTCTCTCAGTGGCTCGACTCTTTCCCATCAAATTTCACTTCTCAGATCAAA 563
QY 816 GCGCAGTATAGTCATATGATGACCTCAGGCAACACCTGGGAGATCCACCATTTGCT 875
DB 564 ATCAGCTCGAATAGATAGTGCACTTTCAAAATATCCACTCAGAAGGAAGAAAGCTGGT 623
QY 876 GCTGAGCAACAAAGCTCCACGGTGGCATGAGCAGATTACGTTGCTGTGCTTAAATTTCCA 935
DB 624 TCTGAAGATATAATCTCTAGTGGCATGAGCACTGCAATGTTGGTGGCTCACTTCCG 683
QY 936 TGGTCAGTCAAGTGGCTTTCGGTTAAGAACTTTTCAAGTTTGGTCAGTGTAGTCACTGTG- 994
DB 684 TGGACGGGTACTGTGTGCTCAGTGAAGAACTTTTCAAGTGGTGGCTCTGATGAAGAAATGG 743
QY 995 --AAGCAGGGCAGACATCTGAGAGGATCATCTCCAGTTTGGGAAGTTGGGAAGGACAT 1052
DB 744 GCCAACTAACAGGAGCAGCAAGGTGATTTCTCCAGTTTGGAAAGATTTGGGAAGGACTT 803

QY 1053 GTTTACCATGGATTATGGATATCCGATTCTCGGTTTCAAGCGTTTGCTATCTGCTGAG 1112
DB 804 GTTCACCATGGACTACCGTTATCCGATATCAGCAATTTCAATCATTTGCAATCTGTCTGAG 863
QY 1113 CAGTTTTGAAACCAAGAAATTTGCCCTGTGAA 1140
DB 864 CAGTTTTGATACCAAAATAGCCTGTGAA 891

Search completed: December 26, 2005, 05:22:21
Job time : 6060 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 23, 2005, 23:20:52 ; Search time 13 Seconds
(without alignments)
208.498 Million cell updates/sec

Title: US-10-763-042-9

Perfect score: 1962

Sequence: 1 WTRSLQEMRSPRHVHA.....SAFQAFCLSSFFETRIACE 380

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 54001

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New:*

- 1: /cgn2_6/prodata/2/pubpaa/US08_NEW_PUB.pap:*
- 2: /cgn2_6/prodata/2/pubpaa/US06_NEW_PUB.pap:*
- 3: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB.pap:*
- 4: /cgn2_6/prodata/2/pubpaa/US09_NEW_PUB.pap:*
- 5: /cgn2_6/prodata/2/pubpaa/US10_NEW_PUB.pap:*
- 6: /cgn2_6/prodata/2/pubpaa/US11_NEW_PUB.pap:*
- 7: /cgn2_6/prodata/2/pubpaa/US12_NEW_PUB.pap:*
- 8: /cgn2_6/prodata/2/pubpaa/US13_NEW_PUB.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	84.5	4.3	865	US-11-109-156-4	Sequence 4, Appli
2	82.5	4.2	761	US-10-485-517-252	Sequence 252, App
3	82	4.2	651	US-11-172-145-10	Sequence 10, Appl
4	80.5	4.1	581	US-11-045-802-30	Sequence 30, Appl
5	80.5	4.1	617	US-11-172-145-8	Sequence 8, Appli
6	80	4.1	483	US-11-137-465-40	Sequence 40, Appl
7	80	4.1	605	US-11-137-465-41	Sequence 41, Appl
8	80	4.1	1075	US-11-174-150-34	Sequence 34, Appl
9	80	4.1	1114	US-11-174-150-35	Sequence 35, Appl
10	79.5	4.1	615	US-11-172-145-6	Sequence 6, Appli
11	79.5	4.1	750	US-11-070-627-2	Sequence 2, Appli
12	79	4.0	584	US-11-045-802-31	Sequence 31, Appl
13	79	4.0	841	US-10-770-726-88	Sequence 88, Appl
14	79	4.0	1897	US-10-821-234-1635	Sequence 1635, Ap
15	79	4.0	1907	US-11-000-463-250	Sequence 250, App
16	77	3.9	521	US-11-091-914-2	Sequence 2, Appli
17	76	3.9	437	US-10-632-150-54	Sequence 54, Appl
18	76	3.9	437	US-11-073-457-54	Sequence 54, Appl
19	76	3.9	437	US-11-073-457-54	Sequence 54, Appl
20	75	3.8	244	US-10-510-386-110	Sequence 110, App
21	75	3.8	826	US-10-821-234-1048	Sequence 1048, Ap
22	74	3.8	397	US-10-467-657-3814	Sequence 3814, Ap
23	73.5	3.7	708	US-11-174-150-25	Sequence 25, Appl
24	73.5	3.7	736	US-11-174-150-26	Sequence 26, Appl
25	73	3.7	415	US-10-467-657-7774	Sequence 7774, Ap

26	72.5	3.7	610	7	US-11-194-246-312	Sequence 312, App
27	72.5	3.7	736	7	US-11-078-189-9	Sequence 9, Appli
28	72.5	3.7	750	7	US-11-070-627-3	Sequence 3, Appli
29	72.5	3.7	1142	7	US-11-044-051-73	Sequence 73, Appl
30	72	3.7	765	6	US-10-821-234-1164	Sequence 1164, Ap
31	72	3.7	934	7	US-11-108-864-19	Sequence 19, Appl
32	72	3.7	945	6	US-10-131-826A-146	Sequence 146, App
33	71.5	3.6	216	7	US-11-055-822-58	Sequence 58, Appl
34	71	3.6	786	7	US-11-070-627-9	Sequence 9, Appli
35	71	3.6	805	6	US-10-467-657-8246	Sequence 8246, Ap
36	70.5	3.6	241	7	US-11-059-275-2	Sequence 2, Appli
37	70.5	3.6	798	7	US-11-107-028-2	Sequence 2, Appli
38	70.5	3.6	881	6	US-10-623-155-430	Sequence 430, App
39	70.5	3.6	920	6	US-10-623-155-357	Sequence 357, App
40	70.5	3.6	943	6	US-10-623-155-161	Sequence 161, App
41	70.5	3.6	1176	6	US-10-821-234-897	Sequence 897, App
42	70	3.6	428	6	US-10-632-150-6	Sequence 6, Appli
43	70	3.6	428	7	US-11-073-457-6	Sequence 6, Appli
44	70	3.6	428	7	US-11-073-460-6	Sequence 6, Appli
45	69.5	3.5	176	6	US-10-467-657-174	Sequence 174, App

ALIGNMENTS

RESULT 1

US-11-109-156-4
; Sequence 4, Application US/11109156
; Publication NO. US20050250144A1
; GENERAL INFORMATION:
; APPLICANT: Toshio Ota
; APPLICANT: Takao Isogai
; APPLICANT: Tetsuo Nishikawa
; APPLICANT: Koji Hayashi
; APPLICANT: Kaoru Otsuka
; APPLICANT: Jun-Ichi Yamamoto
; APPLICANT: Shizuko Ishii
; APPLICANT: Tomoyasu Sugiyama
; APPLICANT: Ai Wakamatsu
; APPLICANT: Keiichi Nagai
; APPLICANT: Tetsuji Otsuki
; APPLICANT: Shin-Ichi Funahashi
; APPLICANT: Chiaki Senoo
; APPLICANT: Jun-Ichi Nezu
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN KINASE/PROTEIN
; TITLE OF INVENTION: PHOSPHATASE
; FILE REFERENCE: 06501-099002
; CURRENT APPLICATION NUMBER: US/11/109,156
; CURRENT FILING DATE: 2005-04-19
; PRIOR APPLICATION NUMBER: US/10/060,065
; PRIOR FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: PCT/JP00/05061
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: US 60/183,322
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: JP 11-248036
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 865
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-109-156-4


```
; Sequence 30, Application US/11045802
; Publication No. US20050257289A1
; GENERAL INFORMATION:
; APPLICANT: Gordon-Kamm, William
; APPLICANT: Helentjaris, Tim
; APPLICANT: Lowe, Keith
; APPLICANT: Shen, Bo
; APPLICANT: Tarczynski, Mitchell
; APPLICANT: Zheng, Peizhong
; TITLE OF INVENTION: AP2 Domain Transcription Factor ODP2 (Ovule Development Protein 2)
; TITLE OF INVENTION: and Methods of Use
; FILE REFERENCE: 035718/286074
; CURRENT APPLICATION NUMBER: US/11/045,802
; CURRENT FILING DATE: 2005-01-28
; PRIOR APPLICATION NUMBER: 60/541,122
; PRIOR FILING DATE: 2004-02-02
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 581
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-11-045-802-30

Query Match      4.1%; Score 80.5; DB 7; Length 581;
Best Local Similarity 19.2%; Pred. No. 2.8;
Matches 61; Conservative 40; Mismatches 87; Indels 129; Gaps 14;

QY 50 TVDGDWPMRRNVVACAGVCRWRILTKRIVAVPEFSSKLTFFISLKQSGPRDSLQV--- 106
DB 171 TSDSNYNNDVVQ-----EKTIVDVETPKKT-----IESFGQRTSIYRGVT 215

QY 107 -----C-----FKRNRNTOSYHL-----YLGTLTSLTDN-- 131
DB 216 RHRWTGRYEHLWDNSCKREGQTRKGGYDKEEKARAYDLAALKYWGTTT--TTNFP 273

QY 132 -KFLLAASKLRATCTDYIISLR--SDDISKESNAY-----LGRMRSN- 172
DB 274 LSEYEKEVEEMKHMTRQEVASURRKSQSGFSRGASIRGVTRHHQHRQWARIGRVAGNK 333

QY 173 --FLGTFTVFDGSGTGAARKQKRSNFIKVSPPVQGSYPIAHISYELNVLGSRGPRR 230
DB 334 DLYLGT-----FTQBERAE-----AYDIAAKF----- 357

QY 231 MRCIMDTPIMSVSGVAVSTSISSFSRSRSPVFRSHSKPLRSNASCSDSGNNIGDPP 290
DB 358 -----RGLSAVTFNFMNRYNVAILESPLFGSSAKRLKDVNNPV--PA 400

QY 291 LVLSNKAPEWHQQLRCW 307
DB 401 MMISNNVSSANNVSGW 417

RESULT 5
US-11-172-145-8
; Sequence 8, Application US/11172145
; Publication No. US20050272696A1
; GENERAL INFORMATION:
; APPLICANT: DeAngelis, Paul
; TITLE OF INVENTION: METHODS OF SELECTIVELY TREATING DISEASES WITH SPECIFIC
; TITLE OF INVENTION: GLYCOSAMINOGLYCAN POLYMERS
; FILE REFERENCE: 3554.104
; CURRENT APPLICATION NUMBER: US/11/172,145
; CURRENT FILING DATE: 2005-06-30
; PRIOR APPLICATION NUMBER: 60/584,442
; PRIOR FILING DATE: 2004-06-30
; PRIOR APPLICATION NUMBER: 10/642,248
; PRIOR FILING DATE: 2003-08-15
; PRIOR APPLICATION NUMBER: 60/404,356
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 60/479,432
; PRIOR FILING DATE: 2003-06-19
; PRIOR APPLICATION NUMBER: 60/491,362
```

```
; PRIOR FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: 10/195,908
; PRIOR FILING DATE: 2002-07-15
; PRIOR APPLICATION NUMBER: 09/437,277
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: 60/107,929
; PRIOR FILING DATE: 1998-11-11
; PRIOR APPLICATION NUMBER: 09/283,402
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/080,414
; PRIOR FILING DATE: 1998-04-02
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 8
; LENGTH: 617
; TYPE: PRT
; ORGANISM: Pasteurella multocida
US-11-172-145-8

Query Match      4.1%; Score 80.5; DB 7; Length 617;
Best Local Similarity 21.4%; Pred. No. 3;
Matches 55; Conservative 30; Mismatches 75; Indels 97; Gaps 14;

QY 7 LOEMSRPRHVVHAAASTANSSDPFSSWSELPEELLREILIRVETVDGGDWPGR-RNVVAC 65
DB 330 LKEIFSFR--IHDALPTISKEMSKLSNPKIP-----VYINICSI-----PSRIKQLQYT 376

QY 66 AGV-----CRSMRILTKRIVAVPEFSSKLTFFISLKQSGPRDSLQVCFIKRNRNTOSYHLY 121
DB 377 IGVLNQCDHPHYLDGYPEVDFIKKL-----GNKATVINC-----QNKNN----- 417

QY 122 LGLTSLTDNGKFLAASKLR-----ATC-----TDYL-----ISL 153
DB 418 ----ESIRDNKGFIKLEKLIKENDGYITCDDDIRYPADYINTWIKKINKYNDKAAIGL 473

QY 154 RSDDISKRSNAYLGRMR-----SNEFLT-----KFTVFDGSGQTGA 188
DB 474 HGVIFPSRVNKFSSDRIVYNFQKPLENDTAVNLTGTVAFRVSIFNKFSLSDFEHGM 533

QY 189 AKWQKS---RSSNFIKV 202
DB 534 VDIYFSILCKKNILQV 550

RESULT 6
US-11-137-465-40
; Sequence 40, Application US/11137465
; Publication No. US2005025558A1
; GENERAL INFORMATION:
; APPLICANT: Agarwal, Pankaj
; APPLICANT: Murdoch, Paul R.
; APPLICANT: Rizvi, Safia, K.
; APPLICANT: Smith, Randall, F.
; APPLICANT: Xiang, Zhaoying
; APPLICANT: Kabnick, Karen
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50018
; CURRENT APPLICATION NUMBER: US/11/137,465
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US/10/239,663
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/09226
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 60/192,158
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,668
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/200,166
; PRIOR FILING DATE: 2000-04-27
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 40
```

```
; LENGTH: 483
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-137-465-40

Query Match          4.1%; Score 80; DB 7; Length 483;
Best Local Similarity 21.9%; Pred. No. 2.4;
Matches 71; Conservative 51; Mismatches 134; Indels 68; Gaps 13;

QY 5 SLLQMRSPHRVVAHAASSTANSSDPFWSSELPEELLREILIRVETVGGDWP--SRRNV 62
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 39 SVAQPLPSLRVADLAANQVMEIFPLTFGEKP--ALRSVYLHNNQLSNAGLPPDAFRGS 96
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 63 VACAGVCSWRILTKIIVAVPEFSSKLTFFPSLKGSGPRDSLVQCFIKENRNTQSYHL 122
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 97 EAIATLSLNNQLSYLPPSLPSLERLHQNLLISKVPRGAL-----SROTQURELYL 149
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 123 GLTTSLTDNGKFLAASKLRATCTDYIISLRSDDISKR-----SNAYLGRM 169
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 150 -QHNQITDSGLDATTFSKLH-----SLEYLDLSHNQLTTPVAGLPTLAILHLGRN 199
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 170 RSNFLGTFTVFDGSGTGAANKQKSRSSNFIKV-----SPRVQGSYPPIAHISYELNVL 223
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 200 RIR-----QVEAARLHGARGRLYLILQHNQLGSSGLPAGALRPLRGLHTLHLY 247
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 224 GS-----RGRPRMRCIMDTIP--MSIVESRGVVASTSISFSSRSPV--FRSHSKP 271
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 248 GNGLDVRVPPALPRRLRALV--LPHNHVAALGARDLVATPGLTELNLAYNRLASARVHHRA 305
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 272 LRSNSA--SCSDSGNNLGDPPPLVL 293
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 306 FRRLRALRSLDLAGNQLTRLPWGL 329
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 7
US-11-137-465-41
; Sequence 41, Application US/11137465
; Publication No. US2005025558A1
; GENERAL INFORMATION:
; APPLICANT: Agarwal, Pankaj
; APPLICANT: Murdoch, Paul R.
; APPLICANT: Rizvi, Safia K.
; APPLICANT: Smith, Randall F.
; APPLICANT: Xiang, Zhaoying
; APPLICANT: Kabnick, Karen
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50018
; CURRENT APPLICATION NUMBER: US/11/137,465
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US/10/239,663
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/09226
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 60/192,158
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,668
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/200,166
; PRIOR FILING DATE: 2000-04-27
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 41
; LENGTH: 605
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-137-465-41

Query Match          4.1%; Score 80; DB 7; Length 605;
Best Local Similarity 21.9%; Pred. No. 3.3;
Matches 71; Conservative 51; Mismatches 134; Indels 68; Gaps 13;

QY 5 SLLQMRSPHRVVAHAASSTANSSDPFWSSELPEELLREILIRVETVGGDWP--SRRNV 62
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 39 SVAQPLPSLRVADLAANQVMEIFPLTFGEKP--ALRSVYLHNNQLSNAGLPPDAFRGS 96
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 63 VACAGVCSWRILTKIIVAVPEFSSKLTFFPSLKGSGPRDSLVQCFIKENRNTQSYHL 122
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 97 EAIATLSLNNQLSYLPPSLPSLERLHQNLLISKVPRGAL-----SROTQURELYL 149
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 123 GLTTSLTDNGKFLAASKLRATCTDYIISLRSDDISKR-----SNAYLGRM 169
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 150 -QHNQITDSGLDATTFSKLH-----SLEYLDLSHNQLTTPVAGLPTLAILHLGRN 199
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 170 RSNFLGTFTVFDGSGTGAANKQKSRSSNFIKV-----SPRVQGSYPPIAHISYELNVL 223
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 200 RIR-----QVEAARLHGARGRLYLILQHNQLGSSGLPAGALRPLRGLHTLHLY 247
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 224 GS-----RGRPRMRCIMDTIP--MSIVESRGVVASTSISFSSRSPV--FRSHSKP 271
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 248 GNGLDVRVPPALPRRLRALV--LPHNHVAALGARDLVATPGLTELNLAYNRLASARVHHRA 305
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 272 LRSNSA--SCSDSGNNLGDPPPLVL 293
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 306 FRRLRALRSLDLAGNQLTRLPWGL 329
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 8
US-11-174-150-34
; Sequence 34, Application US/11174150
; Publication No. US20050260714A1
; GENERAL INFORMATION:
; APPLICANT: Agarwal, Pankaj
; APPLICANT: Murdoch, Paul R.
; APPLICANT: Rizvi, Safia K.
; APPLICANT: Smith, Randall F.
; APPLICANT: Xiang, Zhaoying
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50022
; CURRENT APPLICATION NUMBER: US/11/174,150
; CURRENT FILING DATE: 2005-07-01
; PRIOR APPLICATION NUMBER: US/10/257,174
; PRIOR FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: PCT/US01/11797
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 60/196,603
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/199,417
; PRIOR FILING DATE: 2000-04-24
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 34
; LENGTH: 1075
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-174-150-34

Query Match          4.1%; Score 80; DB 7; Length 1075;
Best Local Similarity 19.7%; Pred. No. 7.3;
Matches 62; Conservative 43; Mismatches 141; Indels 68; Gaps 11;

QY 33 WSELPEELLREILIRVETVGGDWPSSRRNVVACAGVCSWRILTKIIVAVPEFSSKLTTP 92
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 763 WAEGPESAGEPMVVTASTAVAVTVDKRTAIAAAGVQMKLEFLQRFKFWAATRCQSTVDGP 822
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 93 ISLKQSGPRDSLVQCFI-----KENRNTQSYHLVG-----LTTSLTDNGKFLA 137
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 823 CT---QSCEDSLDCFDINDNGFIISKRSRTGR---FLGEVDGAVLTQLLSMGVP--- 873
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 138 ASKLRATCTDYIISLR-----SDDISKRSNAYLGRMRNFLGTFKTFVFDGSGTGAA 189
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 874 ----SQVTWYDYQAMCKPSSHHSSAAQPLVSPISAFLTATRLWLQELVLFLLEWSVWG 929
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 190 KMQKSRSSNFIKVSRRVP--QGSYPPIAHISY-----ELNVLSGRGPRMRCIMDTIPMS-- 241
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 930 YDRGAEAHKKQDPLQPCDTEYFV--FYVQPAIREANGIVECGPCQKVFVVOQIIPNSNL 987
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
QY 242 -----IVESRGVASTSISSFSRSPVFRSHSKPLR-----SNSASC 279
Db 988 LLVTDPTFCRMGSGPEILLTVASAHNASVKCDMRSKQRRRPPDSCHAFHPEENAQDC 1047
QY 280 SDSGNNLGDPLVL 293
Db 1048 GGASDTSASPLLL 1061

RESULT 9
US-11-174-150-35
; Sequence 35, Application US/11174150
; Publication No. US20050260714A1
; GENERAL INFORMATION:
; APPLICANT: Agarwal, Pankaj
; APPLICANT: Murdoch, Paul R.
; APPLICANT: Rizvi, Safia K.
; APPLICANT: Smith, Randall F.
; APPLICANT: Xiang, Zhaoying
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50022
; CURRENT APPLICATION NUMBER: US/11/174,150
; CURRENT FILING DATE: 2005-07-01
; PRIOR APPLICATION NUMBER: US/10/257,174
; PRIOR FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: PCT/US01/11797
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 60/196,603
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/199,417
; PRIOR FILING DATE: 2000-04-24
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 35
; LENGTH: 1114
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-174-150-35

Query Match 4.1%; Score 80; DB 7; Length 1114;
Best Local Similarity 19.7%; Pred. No. 7.6;
Matches 62; Conservative 43; Mismatches 141; Indels 68; Gaps 11;

QY 33 WSELPELRLIRVETVGGWPPRRNVVACAGVCRSWRLITKEIVAVPEPSSKLTTP 92
Db 802 WAEGPESAGEPMVVTASTAVAVTVDKRTAIAAAGVQMKLEFLQRFKFWAATROCSTVDGP 861
QY 93 ISLKQSGPRDSLQCFI-----KRNNTQSYHLVIG-----LTTSLTDNGKFLLA 137
Db 862 CT---QSCSDSDLDCTVIDNNGFILSKRSRGTGR---FLGVDGAVLTQLLSMGVVF--- 912
QY 138 ASKLKRACTDYIISLR-----SDDISKRSNAYLGRMRSNFLTGTFTVFDGSGTGA 189
Db 913 -----SQVTWYDQAMCKPSHHHSAQPLVPSIAFLTATRLQLLVFLLEWSVWGSW 968
QY 190 KMQKRSNFIKVSRRVP-QGSVPIAHISY-----ELNVLSGRGPRMRCIMDTIPMS-- 241
Db 969 YDRGAZAHKHKKQDPLQPCDTEYV--FVYQPAIREANGIVECGPCQKVFVQQIIPNSNL 1026
QY 242 -----IVESRGVASTSISSFSRSPVFRSHSKPLR-----SNSASC 279
Db 1027 LLVTDPTFCRMGSGPEILLTVASAHNASVKCDMRSKQRRRPPDSCHAFHPEENAQDC 1086
QY 280 SDSGNNLGDPLVL 293
Db 1087 GGASDTSASPLLL 1100

RESULT 10
US-11-172-145-6
; Sequence 6, Application US/11172145
; Publication No. US20050272696A1
; GENERAL INFORMATION:
; APPLICANT: DeAngelis, Paul
; TITLE OF INVENTION: METHODS OF SELECTIVELY TREATING DISEASES WITH SPECIFIC
; FILE REFERENCE: 3554.104
; CURRENT APPLICATION NUMBER: US/11/172,145
; CURRENT FILING DATE: 2005-06-30
; PRIOR APPLICATION NUMBER: 60/584,442
; PRIOR FILING DATE: 2004-06-30
; PRIOR APPLICATION NUMBER: 10/642,248
; PRIOR FILING DATE: 2003-08-15
; PRIOR APPLICATION NUMBER: 60/404,356
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 60/479,432
; PRIOR FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: 60/491,362
; PRIOR FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: 10/195,908
; PRIOR FILING DATE: 2002-07-15
; PRIOR APPLICATION NUMBER: 09/437,277
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: 60/107,929
; PRIOR FILING DATE: 1998-11-11
; PRIOR APPLICATION NUMBER: 09/283,402
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/080,414
; PRIOR FILING DATE: 1998-04-02
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 6
; LENGTH: 615
; TYPE: PRT
; ORGANISM: Pasteurella multocida
US-11-172-145-6

Query Match 4.1%; Score 79.5; DB 7; Length 615;
Best Local Similarity 21.0%; Pred. No. 3.8;
Matches 54; Conservative 30; Mismatches 76; Indels 97; Gaps 13;

QY 7 LOEMSRPHRVVHAASSTNSDPFSSWSELPELRLIRVETVGGWPPSR-RNVVAC 65
Db 328 LKEIFSFR--THDALPISKEMSKLSNPKIP-----VVINICSI-----PSRIKQLQYT 374
QY 66 AGV-----CRSWILITKEIVAVPEPSSKLTFTPISLKQSGPRDSLQCFIKRNNTQSHLY 121
Db 375 IGVLNQNCQDHFHYLDGYEVPDFIKLT-----GNKATVINC-----QNK----- 415
QY 122 LGITSLTDNGKFLAASKLR-----ATCTDYI-----ISL 153
Db 416 ----BSIRNGKFLLEKLIKENKQGYITCDDDIRYPADYTNMTIKKYNKAAIGL 471
QY 154 RSDDISKRSNAYLGRMR-----SNFLGT-----KFTVFDGSGTGA 188
Db 472 HGVIPEPRVNYKFSSDRIVYVNFQKPLENDTAVNLTGTVAPRVSIFNKFSLSDFEHPGM 531
QY 189 AKMQKS-----RSSNFIKV 202
Db 532 VDIYFSILCKGNILQV 548

RESULT 11
US-11-070-627-2
; Sequence 2, Application US/11070627
; Publication No. US20050271625A1
; GENERAL INFORMATION:
; APPLICANT: Nash, Kevin R.
; APPLICANT: Burger, Corinna
; TITLE OF INVENTION: RAAV-NEPRILYSIN COMPOSITIONS AND METHODS OF USE
; FILE REFERENCE: 36689.8
; CURRENT APPLICATION NUMBER: US/11/070,627
; CURRENT FILING DATE: 2005-03-02
; PRIOR APPLICATION NUMBER: 60/549,399
; PRIOR FILING DATE: 2004-03-02
```

; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2
; LENGTH: 750
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-11-070-627-2

Query Match 4.1%; Score 79.5; DB 7; Length 750;
Best Local Similarity 25.0%; Pred. No. 4.9; 28; Indels 35; Gaps 4;

Matches 26; Conservative 15; Mismatches 15; Indels 28; Indels 35; Gaps 4;

QY 30 PFSWSLPEELREILIRVETVDGDPKRRNVACAGVCRSWRLTKEIVAVPEFSSKL 89

Db 320 PFSWSNFTWEIMSTVNINQNEE-----EVVVAPEYLTKL 355

QY 90 TFPISLKQSGPRDSLVCQCFIKRNRNTQSYHLVGLTSLTDNGK 133

Db 356 K-PILTKYS-PRDL-----QNLMSWRFIMDLVSSLSRNYK 388

RESULT 12

US-11-045-802-31

; Sequence 31, Application US/11045802

; Publication No. US20050257289A1

; GENERAL INFORMATION:

; APPLICANT: Gordon-Kamm, William

; APPLICANT: Heientjaris, Tim

; APPLICANT: Lowe, Keith

; APPLICANT: Shen, Bo

; APPLICANT: Tarczynski, Mitchell

; APPLICANT: Zheng, Peizhong

; TITLE OF INVENTION: AP2 domain Transcription Factor ODP2 (Ovule Development Protein 2)

; FILE OF INVENTION: and Methods of Use

; FILE REFERENCE: 035718/286074

; CURRENT APPLICATION NUMBER: US/11/045,802

; CURRENT FILING DATE: 2005-01-28

; PRIOR APPLICATION NUMBER: 60/541,122

; PRIOR FILING DATE: 2004-02-02

; NUMBER OF SEQ ID NOS: 38

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 31

; LENGTH: 584

; TYPE: PRT

; ORGANISM: Arabidopsis thaliana

US-11-045-802-31

Query Match 4.0%; Score 79; DB 7; Length 584;

Best Local Similarity 19.4%; Pred. No. 3.9;

Matches 62; Conservative 39; Mismatches 88; Indels 130; Gaps 13;

QY 50 TVDGDWPSRRNVACAGVCRSWRLTKEIVAVPEFSSKLTFPISLKQSGPRDSLVCQCFI 109

Db 171 TSDSNVNNDDVVQ-----EKTIVDVETTPKKT-----IESFGQRTSIYR-GV 214

QY 110 KRNRTQSY--HL-----YLG-----LTTSL 128

Db 215 TRHRVTGRYEAHLWDNSCKEKGQTRGRQVYLGVDYKKEKAARAYDLAALKYWGPTTTN 274

QY 129 TDNGKFLAASKLRATCTDYIISLR--SDDISKRSNAY-----LGRMRS 171

Db 275 FPLSEYEKEVEEMKHTQRYVASLRKSSGFRGASIIYRGVTRHHQHWQARIGRVAG 334

QY 172 N----FLGTFTVPDGGQTGAANKQKRSNFIKVSFVPGQSGYPIAHISYELNVLSGRGP 228

Db 335 NKDLVLTG-----FGTQEEAAE-----AYDIAAIF----- 360

QY 229 RMRRCIMDTIPMSIVSRGVVASTSISFSSRSPPVFRSHSKPLRSNSASCSDSGNNLGD 288

Db 361 -----RGLSAVTNFDNMRYNKAILESPLSPIGSAAKRLKDVNNPV-- 401

QY 289 PPLVLSNKAPRWHEQLRCW 307

Db 402 PAMMISNNVSESANNVSCW 420

RESULT 13

US-10-770-726-88

; Sequence 88, Application US/10770726

; Publication No. US20050266409A1

; GENERAL INFORMATION:

; APPLICANT: Wyeth

; APPLICANT: Brown, Eugene

; APPLICANT: Liu, Wei

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING, PREVENTING, AND TREATING

; FILE OF INVENTION: CANCERS

; FILE REFERENCE: AM101079 (031896-010000)

; CURRENT APPLICATION NUMBER: US/10/770,726

; CURRENT FILING DATE: 2004-02-04

; NUMBER OF SEQ ID NOS: 48640

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 88

; LENGTH: 841

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-770-726-88

Query Match 4.0%; Score 79; DB 6; Length 841;

Best Local Similarity 19.6%; Pred. No. 6.5;

Matches 90; Conservative 62; Mismatches 175; Indels 132; Gaps 21;

QY 4 RSLQEMRSR---PHRVVHAAASTANSSDPFWSLPEELL--REILIRVETVDG--GDM 56

Db 147 KQLQKAVERGAVPLEMLEIALNLNLQKQLLSEEEKNLASATVLTAEQSFSGSLGHL 206

QY 57 PSRRNVACAGVCRSWRLTKEIVAVPEFSSKLTFPISL-----KQSGP----- 100

Db 207 QNRNNSCDSRGQTTKARELYGE--NMPPQDAEIGYRNSLRQTNKTKQSCFGRVVPVNLN 264

QY 101 -----RDSLVCQCFIKRN-----RNTOSYHLYLGLTSL 128

Db 265 SPDCDVKTDDSVVPCFMKQTSRSCRDLVPGSKPNSGNDSCELNLS----- 313

QY 129 TDNGKF---LLAASKLRATCTDYIISLRSDDISKRSNAYLGRMRSNFIKGTFTVPDGSQ 185

Db 314 VQNSHFKEPLVSDKSELIITD-SITLK---NKTESLLAKLEET---KEYQPEVEPE 365

QY 186 TCAAKMQKRSNFIKVSFVPGQSGYPIAHISYELNVLSGRGPRMRCIMDTIPMSIVS 245

Db 366 SNQKQWQSKRSECIQNQNPAASSNHWQIPELARKVN-----TEQ 404

QY 246 RGWVASTSISFSSRSPPVFRSHSKPLRSNSASCSDSGNNLGD-----PPLVLSNKAPR 299

Db 405 KHTTEQPVFSVSKQSPFI--STSKWFDPKSICKTPSSNTLDDYMSCFRTPVVKNDFFPA 462

QY 300 W-----HEQLRCWCLNFHGRVTVASVKNFQLVAVSDC--EAGQTSR---IILOFGKVG-- 348

Db 463 CQLSTPYGPQAFQOQOH-QILATPLQNLQVLASSANECISVKGRIYSILKQIGSGSS 521

QY 349 -----KMFMTWDYGPISAFQAFALCISSEFETRIA 378

Db 522 KVFQVLNEKKQIYAIKY---VNLSEADNQTLDYSYRNEIA 557

RESULT 14

US-10-821-234-1635

; Sequence 1635, Application US/10821234

; Publication No. US20050255114A1

; GENERAL INFORMATION:

; APPLICANT: Labat, Ivan

; APPLICANT: Stache-Crain, Birgit

; APPLICANT: Andarmani, Susan

; APPLICANT: Tang, Y. Tom

; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia

; FILE REFERENCE: 821A

; CURRENT APPLICATION NUMBER: US/10/821,234

; CURRENT FILING DATE: 2004-04-07
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_seq_genes Version 1.0
; SEQ ID NO 1635
; LENGTH: 1897
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1635

Query Match 4.0%; Score 79; DB 6; Length 1897;
Best Local Similarity 23.8%; Pred. No. 20;
Matches 85; Conservative 38; Mismatches 138; Indels 96; Gaps 19;
QY 14 PHRVHAAASTANSSDPFSW---SELPEELL-----REILRVETVGGDWPSRRNVV 63
DB 811 PGRPTMMISTAMNTALLQWHPKELPGELLYGRLQYCRADARPNTIDFGKDDQHTVT 870
QY 64 AC-AGVCRWRILITKEIVAV-PFSSKLTFFPSLKQSGPRDSLVCQFIKRNRTQSYHLY 121
DB 871 GLHKGTTYIFRLAAKNRAGLGEFEKEIIRTPEDLPSGFP-----QNLHV- 914
QY 122 LGTTTSLTD-----NGKFL-----AASKLKRATC-----TDY 149
DB 915 TGLTTTTELAWDPPVLAERNGRIISYTVVFRDINSQQLQNTTDTTRFTLTGLKPDTTY 974
QY 150 IISLRDDDISKRSNAYLGRMRSNFLTGTFTVFGSOTGAQKQKSSNFIVKSPRPQG 209
DB 975 DIKVRW-WTSKSGPLSPSIQSRTPVE-QVFAKNFRVAAAMKTS-----VLLSWEVPD- 1026
QY 210 SYPIA---HISYELNVLGSRGPRRCIMDTIPMS-----IVESRGVVASTSISFSSRS 262
DB 1027 SYKSAVFPFKILYNGQSVEDVGHSMRKLADLQNTTEYSFVLMNRGSSAGGLQHLVIRTA 1086
QY 263 PVFRSHSKPLRNSASCSGSDGNNLG-----DPLVL-----SNKAPRW 300
DB 1087 PDLPLH-KPLPA-SAYIEDGRFDSMPHVQDPSLVRWFYIVVPIRDRVGSMLTPRW 1141

RESULT 15
US-11-000-463-250
; Sequence 250, Application US/11000463
; Publication No. US20050266423A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Chen, Rui-hong B.
; APPLICANT: Qian, Xiaohong B.
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Zhou, Ping
; APPLICANT: Cao, Yi-Cheng
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 785GIP4CN
; CURRENT FILING DATE: 2004-11-29
; PRIOR FILING DATE: 2002-11-08
; PRIOR FILING DATE: 2002-11-08
; PRIOR FILING DATE: 2001-01-25
; PRIOR FILING DATE: 2001-01-25
; PRIOR FILING DATE: 2001-08-03
; PRIOR FILING DATE: 2001-08-03
; PRIOR FILING DATE: 2000-01-25
; PRIOR FILING DATE: 2000-07-17
; PRIOR FILING DATE: 2000-07-17
; PRIOR FILING DATE: 2000-08-03
; PRIOR FILING DATE: 2000-08-03
; PRIOR FILING DATE: 2000-08-03

; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 944
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 250
; LENGTH: 1907
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-000-463-250

Query Match 4.0%; Score 79; DB 7; Length 1907;
Best Local Similarity 23.8%; Pred. No. 20;
Matches 85; Conservative 38; Mismatches 138; Indels 96; Gaps 19;
QY 14 PHRVHAAASTANSSDPFSW---SELPEELL-----REILRVETVGGDWPSRRNVV 63
DB 821 PGRPTMMISTAMNTALLQWHPKELPGELLYGRLQYCRADARPNTIDFGKDDQHTVT 880
QY 64 AC-AGVCRWRILITKEIVAV-PFSSKLTFFPSLKQSGPRDSLVCQFIKRNRTQSYHLY 121
DB 881 GLHKGTTYIFRLAAKNRAGLGEFEKEIIRTPEDLPSGFP-----QNLHV- 924
QY 122 LGTTTSLTD-----NGKFL-----AASKLKRATC-----TDY 149
DB 925 TGLTTTTELAWDPPVLAERNGRIISYTVVFRDINSQQLQNTTDTTRFTLTGLKPDTTY 984
QY 150 IISLRDDDISKRSNAYLGRMRSNFLTGTFTVFGSOTGAQKQKSSNFIVKSPRPQG 209
DB 985 DIKVRW-WTSKSGPLSPSIQSRTPVE-QVFAKNFRVAAAMKTS-----VLLSWEVPD- 1036
QY 210 SYPIA---HISYELNVLGSRGPRRCIMDTIPMS-----IVESRGVVASTSISFSSRS 262
DB 1037 SYKSAVFPFKILYNGQSVEDVGHSMRKLADLQNTTEYSFVLMNRGSSAGGLQHLVIRTA 1096
QY 263 PVFRSHSKPLRNSASCSGSDGNNLG-----DPLVL-----SNKAPRW 300
DB 1097 PDLPLH-KPLPA-SAYIEDGRFDSMPHVQDPSLVRWFYIVVPIRDRVGSMLTPRW 1151

Search completed: December 23, 2005, 23:29:44
Job time : 17 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 23, 2005, 22:46:55 ; Search time 198 Seconds
(without alignments)
843.252 Million cell updates/sec

Title: US-10-763-042-9
Perfect score: 1962
Sequence: 1 MTRFLLQEMSRPHRVHA.....SAFOAFAICLSSFETRIACE 380

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A: Geneseq 21: *
1: geneseqp1980s: *
2: geneseqp1990s: *
3: geneseqp2000s: *
4: geneseqp2001s: *
5: geneseqp2002s: *
6: geneseqp2003as: *
7: geneseqp2003bs: *
8: geneseqp2004s: *
9: geneseqp2005s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1962	100.0	380	8	Adn73049 Thale cre
2	1962	100.0	380	9	Adw64863 Arabidops
3	1442.5	73.5	389	8	Ado61995 Transcrip
4	1437.5	73.3	389	3	Aag15637 Arabidops
5	1414.5	72.1	380	9	Adw64865 Arabidops
6	1409.5	71.8	380	3	Aag15638 Arabidops
7	1215.5	62.0	403	9	Adw17479 Eucalyptu
8	1153.5	58.8	408	9	Adw17480 Eucalyptu
9	1149	58.6	406	8	Adn73405 Thale cre
10	1149	58.6	406	9	Adw64857 Arabidops
11	1142.5	58.2	308	3	Aag15639 Arabidops
12	1119.5	57.1	411	9	Adw17849 Pinus rad
13	1082	55.1	418	9	Adw17851 Pinus rad
14	1049.5	53.5	418	9	Adw17850 Pinus rad
15	1028	52.4	414	9	Adw18506 Pinus rad
16	1028	52.4	414	9	Adw18507 Pinus rad
17	1018	51.9	436	8	Ady24685 plant ful
18	1018	51.9	436	8	Adx97136 plant ful
19	1013.5	51.7	455	9	Adw64907 Arabidops
20	1010.5	51.5	437	8	Adx94433 plant ful
21	1006	51.3	442	8	Adx97135 plant ful
22	1002	51.1	412	9	Adw17848 Pinus rad
23	999.5	50.9	445	7	Adx37191 plant yie
24	999.5	50.9	445	8	Adi44433 Plant tra

ALIGNMENTS

RESULT 1
ADN73049
ID ADN73049 standard; protein; 380 AA.
XX
AC ADN73049;
XX
DT 15-JUL-2004 (first entry)
XX
DE Thale cress protein upregulated in E2Fa/Dpa expressing plants SeqID 944.
XX
KW plant; transgenic; E2Fa/Dpa transcription factor; growth regulator;
KW animal feed product; thale cress; cell wall biosynthesis;
KW nitrogen metabolism; carbon metabolism.
XX
OS Arabidopsis thaliana.
XX
PN WO2004035798-A2.
XX
PD 29-APR-2004.
XX
PF 20-OCT-2003; 2003WO-EP011658.
XX
PR 18-OCT-2002; 2002EP-00079408.
XX
(CROP-) CROPDESIGN NV.
XX
PI Inze D, De Veylder L, Vlieghe K;
XX
WPI; 2004-348466/32.
DR N-PSDB; ADN73048.

Altering plant characteristics, useful for producing plants for enzyme or pharmaceutical production comprises modifying in a plant, expression of one or more nucleic acids and/or modifying level or activity of one or more proteins.
Claim 1; SEQ ID NO 944; 134pp; English.

This invention relates to a novel method for altering one or more plant characteristics. Specifically, it refers to identifying genes that are up- or down-regulated in transgenic plants overexpressing the heterodimeric E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to alter plant characteristics accordingly. The present invention describes generating transgenic plants for the production of growth regulators, enzymes, therapeutics, pharmaceuticals and animal feed products, where the altered plant characteristics are selected from increased yield or biomass, enhanced survival capacity, stress tolerance, plant architecture

25 999.5 50.9 445 9 ADW64864 Arabidops
26 991.5 50.5 450 9 ADW64855 Arabidops
27 989 50.4 468 8 ADI06759 Plant ful
28 987 50.3 424 9 ADW17484 Eucalyptu
29 987 50.3 424 9 ADW17483 Eucalyptu
30 987 50.3 424 9 ADW17485 Eucalyptu
31 982 50.1 415 9 ADW17481 Eucalyptu
32 980 49.9 408 9 ADW17847 Pinus rad
33 972 49.5 527 8 ADX97137 plant ful
34 971 49.5 551 8 ADY09712 plant ful
35 966.5 49.3 394 9 ADW64856 Arabidops
36 955 48.7 480 8 ADW48240 Polypepti
37 950.5 48.4 429 8 ADW61991 Transcrip
38 950.5 48.4 429 9 ADW64859 Arabidops
39 929 47.3 388 9 ADW64860 Arabidops
40 858.5 43.8 421 9 ADW17482 Eucalyptu
41 843 43.0 379 3 AAG07737 Arabidops
42 843 43.0 379 9 ADW64861 Arabidops
43 843 43.0 400 3 AAG07736 Arabidops
44 831.5 42.4 370 8 ADT58092 Plant pol
45 830.5 42.3 342 3 AAG07738 Arabidops

CC or physiology, altered endoreduplication, biochemistry, signal
CC transduction, storage lipid mobilization and/or altered photosynthesis,
CC each relative to the corresponding wild type plants. Accordingly, these
CC sequences can also be useful as positive or negative selectable markers
CC during transformation of cells or tissues. The identified genes play a
CC role in a variety of biological processes such as DNA replication, cell
CC wall biosynthesis, nitrogen and/or carbon metabolism or they function as
CC transcription factors. This polypeptide sequence is thale cress protein
CC expressed by a gene upregulated 1.3 fold or more in plants overexpressing
CC the E2Fa/Dfa transcription factor, given in an exemplification of the
CC invention.
XX
XX
SQ Sequence 380 AA;

Query Match 100.0%; Score 1962; DB 8; Length 380;
Best Local Similarity 100.0%; Pred. No. 5.4e-211;
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTFRSLQEMRSRPHRVVHAAASTANSSDPFWSSELPEELLREILIRVETVDGGWPSPRR 60
DB 1 MTFRSLQEMRSRPHRVVHAAASTANSSDPFWSSELPEELLREILIRVETVDGGWPSPRR 60
QY 61 NVVACAGVCRSWRIITKEIVAVPEFSSKLTFFPISLKQSGPRDSLVOCFIKRNRNTQSYHL 120
DB 61 NVVACAGVCRSWRIITKEIVAVPEFSSKLTFFPISLKQSGPRDSLVOCFIKRNRNTQSYHL 120
QY 121 YLGTTSLTDNGKFLAASKLKRATCTDYIISLRSDDISKRSNAYLGRMRSNFLTGTFTV 180
DB 121 YLGTTSLTDNGKFLAASKLKRATCTDYIISLRSDDISKRSNAYLGRMRSNFLTGTFTV 180
QY 181 FDGSOTGAAMQKRSNNFIKVSPPVQGSYPYIAHISYELNVLGSRGPRRMCIMDTIPM 240
DB 181 FDGSOTGAAMQKRSNNFIKVSPPVQGSYPYIAHISYELNVLGSRGPRRMCIMDTIPM 240
QY 241 SIVESRGVAVASTISSFSRSRSPVFRSHSKPLRSNSASCSDSGNLGDPLVLSNKAPRW 300
DB 241 SIVESRGVAVASTISSFSRSRSPVFRSHSKPLRSNSASCSDSGNLGDPLVLSNKAPRW 300
QY 301 HEQLRCWCLNFHGRVTVASVKNFQLVAVSDCEAGQTSERIILQFGKVGKDMFTMDYGYPI 360
DB 301 HEQLRCWCLNFHGRVTVASVKNFQLVAVSDCEAGQTSERIILQFGKVGKDMFTMDYGYPI 360
QY 361 SAFOAFAICLSSPETRIACE 380
DB 361 SAFOAFAICLSSPETRIACE 380

RESULT 2
ADM64863
ID ADM64863 standard; protein; 380 AA.

XX ADM64863;
XX
XX
XX 07-APR-2005 (first entry)
XX Arabidopsis thaliana TUBBY-like protein (TLP) 9.
XX
XX Transgenic plant; salt tolerance; crop improvement; cold tolerance;
KW oxidative stress; drought resistance; TUBBY-like protein; plant.
XX Arabidopsis thaliana.
XX
XX US2005014266-A1.
XX

PD 20-JAN-2005.
XX
XX 21-JAN-2004; 2004US-00763042.
XX
XX 21-JAN-2003; 2003US-0441380P.
PR (SINT-) ACAD SINICA.
XX
XX Shaw J, Lai C;

XX WPI; 2005-080955/09.
DR N-PSDB; ADM64874.
XX
PT New isolated Arabidopsis TUBBY-like proteins, useful for producing
PT transgenic or transformed cells or plants having higher tolerance to
PT salt, chilling, pathogens, oxidative stress, or water-deficit.
XX
PS Claim 1; SEQ ID NO 9; 43pp; English.
XX
CC The invention relates to Arabidopsis TUBBY-like proteins (TLP) and their
CC corresponding nucleic acid sequences. The invention also relates to a
CC transformed cell or a transgenic plant containing TLP nucleic acid and a
CC transformed cell or a transgenic plant lacking TLP nucleic acid. The TLP
CC polypeptides and nucleic acids are useful for producing transgenic or
CC transformed cells or plants having higher tolerance to salt, chilling, the
CC pathogens, oxidative stress or water-deficit. The present sequence is the
CC Arabidopsis thaliana TLP 9 protein.
XX
SQ Sequence 380 AA;

Query Match 100.0%; Score 1962; DB 9; Length 380;
Best Local Similarity 100.0%; Pred. No. 5.4e-211;
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTFRSLQEMRSRPHRVVHAAASTANSSDPFWSSELPEELLREILIRVETVDGGWPSPRR 60
DB 1 MTFRSLQEMRSRPHRVVHAAASTANSSDPFWSSELPEELLREILIRVETVDGGWPSPRR 60
QY 61 NVVACAGVCRSWRIITKEIVAVPEFSSKLTFFPISLKQSGPRDSLVOCFIKRNRNTQSYHL 120
DB 61 NVVACAGVCRSWRIITKEIVAVPEFSSKLTFFPISLKQSGPRDSLVOCFIKRNRNTQSYHL 120
QY 121 YLGTTSLTDNGKFLAASKLKRATCTDYIISLRSDDISKRSNAYLGRMRSNFLTGTFTV 180
DB 121 YLGTTSLTDNGKFLAASKLKRATCTDYIISLRSDDISKRSNAYLGRMRSNFLTGTFTV 180
QY 181 FDGSOTGAAMQKRSNNFIKVSPPVQGSYPYIAHISYELNVLGSRGPRRMCIMDTIPM 240
DB 181 FDGSOTGAAMQKRSNNFIKVSPPVQGSYPYIAHISYELNVLGSRGPRRMCIMDTIPM 240
QY 241 SIVESRGVAVASTISSFSRSRSPVFRSHSKPLRSNSASCSDSGNLGDPLVLSNKAPRW 300
DB 241 SIVESRGVAVASTISSFSRSRSPVFRSHSKPLRSNSASCSDSGNLGDPLVLSNKAPRW 300
QY 301 HEQLRCWCLNFHGRVTVASVKNFQLVAVSDCEAGQTSERIILQFGKVGKDMFTMDYGYPI 360
DB 301 HEQLRCWCLNFHGRVTVASVKNFQLVAVSDCEAGQTSERIILQFGKVGKDMFTMDYGYPI 360
QY 361 SAFOAFAICLSSPETRIACE 380
DB 361 SAFOAFAICLSSPETRIACE 380

RESULT 3
ADO61995
ID ADO61995 standard; protein; 389 AA.

XX ADO61995;
XX
XX 15-JUL-2004 (first entry)
XX
XX Transcription factor G2603, SEQ ID 462.
XX
XX Plant; transcription factor; transgenic plant; abiotic stress tolerance;
KW osmotic stress tolerance; cold tolerance; heat tolerance;
KW low nitrogen tolerance; low phosphate tolerance; fungal disease;
KW glyphosate resistance; flowering; fertility; seed development.
XX Arabidopsis thaliana.
XX
XX WO2004031349-A2.
XX

PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 20-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144632P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.

PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155133P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159291P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.

```

PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161320P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 73.3%; Score 1437.5; DB 3; Length 389;
Best Local Similarity 72.5%; Pred. No. 5.7e-152; Indels 23; Gaps 5;
Matches 287; Conservative 32; Mismatches 54;

QY 1 MTPRSLLOEMRSPHRVH-----AAASTANSDDPFSWSELPEELLREILIRVETVDGG 54
DB 1 MTLRSLILEMRSPHRVHDLAAAAADSTSVSSQDYRWSEIPEELLREILIRVEADGG 60
QY 55 DWPSRRNVACAGVCSRWILTKGIIVAVPEFSSKLTFFISLKQSGPRDSLVCQFIKRN 114
DB 61 GWPSSRRNVACAGVCRGWRLLMNETVVVPEISSKLTFFISLKQSGPRDSLVCQFIKRN 120
QY 115 TQSYHLYLGLTSLTDNGKFLAASKLKRATCTDYIISLRSDDISKRSNAYLGRMSN 174
DB 121 TQSYHLYLGLTSLTDNGKFLAASKLKRATCTDYIISLRSDDISKRSNAYLGRMSN 180
QY 175 GYKFTVFDGS---QTGAAMKOKSRSSNFIKVSPRVPQGSYPYTAHISYELNVLGSRGPRM 231
DB 181 GYKFTVFDGNLPLSTGAALKRSRYPNPAKSAKVLGSPYVAHYTYELNVLGSRGPRM 240
QY 232 RCIMDTIPMSIVESRGVASTSISFSSRSSPVPFRSHSKPLRSNSASCSDSGNNLGDPL 291
DB 241 QCLMDIPTSTWBPQGVASEPSEFPLLGTRSTLSRSQSKPLRSS-----SHLKETPL 293
QY 292 VLSNKAAPRHEQRCWCLNPHGRVTVASVKNFOLVAV-SDCEAG-----QTSERIILOF 344
DB 294 VLSNKAAPRHEQRCWCLNPHGRVTVASVKNFOLVAV-SDCEAG-----QTSERIILOF 353
QY 345 GKVGKDMFTMDYGPISAFQAFALCLSSFETRIACE 380
DB 354 GKVGKDMFTMDYGPISAFQAFALCLSSFETRIACE 389

RESULT 5
ADM64865
ID ADM64865 standard; protein; 380 AA.
XX
AC ADM64865;
XX
DT 07-APR-2005 (first entry)
XX
DE Arabidopsis thaliana TUBBY-like protein (TLP) 11.
XX
KW Transgenic plant; salt tolerance; crop improvement; cold tolerance;
KW oxidative stress; drought resistance; TUBBY-like protein; plant.
XX
OS Arabidopsis thaliana.
XX
PN US2005014266-A1.
XX
PD 20-JAN-2005.
XX
PF 21-JAN-2004; 2004US-00763042.
XX
PR 21-JAN-2003; 2003US-0441380P.
XX
PA (SINI-) ACAD SINICA.
XX
PI Shaw J, Lai C;
XX
XX WPI; 2005-080955/09.
DR N-PSDB; ADM64876.
XX
XX New isolated Arabidopsis TUBBY-like proteins, useful for producing
PT transgenic or transformed cells or plants having higher tolerance to
PT salt, chilling, pathogens, oxidative stress, or water-deficit.
XX

```

```

PS Claim 1; SEQ ID NO 11; 43pp; English.
XX
CC The invention relates to Arabidopsis TUBBY-like proteins (TLP) and their
CC corresponding nucleic acid sequences. The invention also relates to a
CC transformed cell or a transgenic plant containing TLP nucleic acid and a
CC transpeptidase and nucleic acids are useful for producing transgenic or
CC transformed cells or plants having higher tolerance to salt, chilling, the
CC pathogens, oxidative stress or water-deficit. The present sequence is the
XX Arabidopsis thaliana TLP 11 protein.
SQ Sequence 380 AA;

Query Match 72.1%; Score 1414.5; DB 9; Length 380;
Best Local Similarity 72.9%; Pred. No. 2.1e-149; Indels 23; Gaps 5;
Matches 282; Conservative 31; Mismatches 51;

QY 10 MRSRPHRVH-----AAASTANSDDPFSWSELPEELLREILIRVETVDGGWPSRRNV 63
DB 1 MRSRPHRVHDLAAAAADSTSVSSQDYRWSEIPEELLREILIRVEADGGWPSRRNV 60
QY 64 ACAGVCSRWILTKGIIVAVPEFSSKLTFFISLKQSGPRDSLVCQFIKRNNTQSYHLYG 123
DB 61 ACAGVCSRWILTKGIIVAVPEFSSKLTFFISLKQSGPRDSLVCQFIKRNNTQSYHLYG 120
QY 124 LTSLTNDNGKFLAASKLKRATCTDYIISLRSDDISKRSNAYLGRMSNPLGKFTVFDG 183
DB 121 LTSLTNDNGKFLAASKLKRATCTDYIISLRSDDISKRSNAYLGRMSNPLGKFTVFDG 180
QY 184 S---QTGAAMKOKSRSSNFIKVSPRVPQGSYPYTAHISYELNVLGSRGPRMRCIMDTIP 240
DB 181 NLLPSTGAALKRSRYPNPAKSAKVLGSPYVAHYTYELNVLGSRGPRMRCIMDTIP 240
QY 241 SIVESRGVASTSISFSSRSSPVPFRSHSKPLRSNSASCSDSGNNLGDPLVLSNKAAPRW 300
DB 241 STWEPQGVASEPSEFPLLGTRSTLSRSQSKPLRSS-----SHLKETPLVLSNKAAPRW 293
QY 301 HEQLRCWCLNPHGRVTVASVKNFOLVAV-SDCEAG-----QTSERIILOFGKVGKDMFT 353
DB 294 HEQLRCWCLNPHGRVTVASVKNFOLVAV-SDCEAG-----QTSERIILOFGKVGKDMFT 353
QY 354 MDYGPISAFQAFALCLSSFETRIACE 380
DB 354 MDYGPISAFQAFALCLSSFETRIACE 380

RESULT 6
AAG15638
ID AAG15638 standard; protein; 380 AA.
XX
AC AAG15638;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 15969.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.

```

PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 04-MAY-1999; 99US-0132048P.
PR 05-MAY-1999; 99US-0132407P.
PR 06-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 18-MAY-1999; 99US-0134370P.
PR 19-MAY-1999; 99US-0134768P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143342P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143342P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 31-AUG-1999; 99US-0151303P.
PR 01-SEP-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 10-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 24-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.

QY	275	NSASCSDSGNLGG--DPPLVLSNKAPRWHEQLRCWCLNFHGRVTVASVKNFOLVAVSDCE	332	Db	3	SFKSIIDVDKDBIGNISRKGFHVKFGYGLASRSRSHRVVDSSVVVDALKQSCWASMPPEL
Db	295	NRAESVSSPLVSQKDGALVLRNKAPRWHEQLRCWCLNFHGRVTVASVKNFOLVAVSPENE	354	QY	41	LREILIRVETDGDWPSRRNVACAGYCRSRWRIILTKEIVAVPESSKLTFFISLKQSGP
QY	333	-AGOTSERIIQFGKVGKDMFTMDYGPISAFQAFALCLSSPETRIACE	380	Db	63	LRDVLMRIKSE-STWPPRKNVACAGYCRNRWREIMKEIVKTPFVSGKLTFFISLKQSGP
Db	355	PAGPQHEKIIQFGKVGKDLFTMDYRYPISAFQAFALCLSSPDTKIACE	403	QY	101	RDSLVOQCFIKENRNTQSVHLYLGLTSLTDNGKFLAASKLKRATCTDYIISLRSDDISK
RESULT 8				Db	122	RDSLQCYIKRNRNQTYLYLSLQALNEDGKFLAARKLRPTCTDYVLSLNPDDMSK
ADW17480				QY	161	RSNAYLGRMSNFLTGFVFDGSGTGA-AKMOKSRSS---NFIKVSRRVPOGSGYPIAH
ID	ADW17480	standard; protein; 408 AA.		Db	182	GSSTYIGKLRSNFLTGFVFDGSGTGA-AKMOKSRSS---NFIKVSRRVPOGSGYPIAH
AC	ADW17480;			QY	217	SYELNVLSGRGPRRMRCTMDIPMSIVESRGVV-----ASTSISFSRSRSPFRSHSK
XX	24-MAR-2005	(first entry)		Db	242	SYELNVLSGRGPRRMRCTMDIPMSIVESRGVV-----ASTSISFSRSRSPFRSHSK
DE	Eucalyptus grandis	transcription factor protein TUBBY family seq 1242.		QY	271	PLRSNSASCSDSGNLGDPP---LVLNKAPEWHQRLRCWCLNFHGRVTVASVKNFOLVA
XX	plant; transcription; gene regulation; gene expression; transgenic plant;			Db	297	--KSTRVENLESGRPMSAPKQGLVLRNKSRAWHQRLRCWCLNFHGRVTVASVKNFOLVA
KW	drought resistance; disease resistance; salt tolerance; cold tolerance;			QY	328	VSD-CEAGOTSERIIQFGKVGKDMFTMDYGPISAFQAFALCLSSPETRIACE
KW	freezing tolerance; flowering; flavor enhancer; flower color.			Db	355	SPEQGVAGTEHENVILQFGKVGKDVFTMDYQYPISAFQSFALCLSSFDTRIACE
XX	Eucalyptus grandis.			RESULT 9		
XX	WO2005001050-A2.			ADW173405		
PN	06-JAN-2005.			ID	ADW173405	standard; protein; 406 AA.
XX	07-JUN-2004;	2004WO-US017965.		XX	AC	ADW173405;
XX	06-JUN-2003;	2003US-0476189P.		XX	DT	15-JUL-2004 (first entry)
PA	(ARBO-) ARBORGEN LLC.			XX	DE	Thale cress protein upregulated in E2Fa/Dpa expressing plants seqID 1300.
XX	Blokeberg LM, Bryant C, Connert MB, Emerson SJ, Frost MJ;			XX	KW	plant; transgenic; E2Fa/Dpa transcription factor; growth regulator;
PI	Forster RLS, Grigor M, Higgins C, Laeham A, Lund ST, Magnus A;			KW	XX	animal feed product; thale cress; cell wall biosynthesis;
PI	Phillips J, Puthigae S, Veerakone S, Westwood C, Gause K, Wood M;			XX	XX	nitrogen metabolism; carbon metabolism.
DR	WPI; 2005-075542/08.			OS	OS	Arabidopsis thaliana.
DR	N-PSDB; ADW16694.			XX	PN	WO2004035798-A2.
XX	New polynucleotides isolated from plants encoding transcription factors,			XX	PD	29-APR-2004.
PT	and polypeptides encoded by such polynucleotides, useful for regulating			XX	PF	20-OCT-2003; 2003WO-EP011658.
PT	gene transcription and gene expression.			XX	PR	18-OCT-2002; 2002EP-00079408.
XX	Claim 31; SEQ ID NO 1242; 1265pp; English.			XX	PA	(CROP-) CROPEDESIGN NV.
XX	This invention relates to novel isolated plant nucleic acid molecules, or			XX	PI	Inze D, De Veylder L, Vlieghe K;
CC	variants thereof, that encode transcription factors. Specifically, it			XX	XX	WPI; 2004-348466/32.
CC	refers to transcription factor proteins that are capable of binding to			DR	DR	N-PSDB; ADW173404.
CC	DNA in order to regulate gene transcription and gene expression in a			XX	PT	Altering plant characteristics, useful for producing plants for enzyme or
CC	plants, in particular Eucalyptus grandis and Pinus radiata. The present			PT	PT	pharmaceutical production comprises modifying in a plant, expression of
CC	invention describes DNA constructs containing DNA encoding a			PT	PT	one or more nucleic acids and/or modifying level or activity of one or
CC	transcription factor that regulates the promoter, which is operably			XX	XX	more proteins.
CC	linked to the desired nucleic acid to be expressed. It further provides			XX	PS	Claim 1; SEQ ID NO 1300; 134pp; English.
CC	transgenic plants expressing a transcription factor that confers a trait			XX	CC	This invention relates to a novel method for altering one or more plant
CC	to the plant such as increased drought, salt or disease tolerance, height			CC	CC	characteristics. Specifically, it refers to identifying genes that are up
CC	change, enhanced cold/frost tolerance, enhanced color, health and			CC	CC	or down-regulated in transgenic plants overexpressing the heterodimeric
CC	nutritional characteristics, as well as improved taste, starch			CC	CC	E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to
CC	composition, flower longevity and germination, amongst others.			CC	CC	alter plant characteristics accordingly. The present invention describes
CC	Accordingly, such plants that are successfully transfected with a DNA			CC	CC	generating transgenic plants for the production of growth regulators,
CC	construct can be characterized by a difference in flower color, petal or			CC	CC	enzymes, therapeutics, pharmaceuticals and animal feed products, where
CC	leaf shape and size, aroma or plant height. This polypeptide is a plant					
XX	transcription factor protein sequence of the invention.					
XX	Sequence 408 AA;					
SQ	Query Match	58.8%; Score 1153.5; DB 9; Length 408;				
	Best Local Similarity	56.0%; Pred. No. 5.5e-120;				
	Matches 233; Conservative	60; Mismatches 79; Indels 43; Gaps 9;				
QY	2	TFRSLQEQ-----MRSRPHRVVHAASSTANSSDPFSNSLPEEL 40				

XX 06-SEP-2000. 99US-0142055P.
PD 06-JUL-1999; 99US-0142390P.
XX 08-JUL-1999; 99US-0142803P.
PF 09-JUL-1999; 99US-0142920P.
XX 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0144005P.
PR 15-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 12-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 16-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 07-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.

	Db	180	RGSNRYIGKLRNFNGTGYTDYDQDPCTGAIASTSRGSRVRVSGSKQVSRRVPAGSVVAH	239
	Qy	216	ISVELNVLSRGPRMRCTMDTPMWSIVESRGGVVASTSISFSRSRSPVFRSH----	SXP 271
	Db	240	IAYELNVLTGRPPRMQCIMDAVASAMEACGSAPTTECPLESLDSLATLPHGLKSEV	299
	Qy	272	LRSNASCSDSGNNLGDPLVLNSNKAPRWHEBQLRCWCLFNHGRTVTASVKNFOLVAUSD-	330
	Db	300	IETSFPSGPINSNWKEGPLLLKNKAPRWHELQWCWLKFNGRVTVPSVKNFOLVAATEP	359
	Qy	331	--CEAQQTSERIILOFGVGKDMMFTWDYGYPISAFQAFAICLSFETRIACE	380
	Db	360	TOPAAOSDHDKVILOFKGVKGOTFMTRYLPFLSAFOAFAICLSFDFTKFACE	411

RESULT 13	
ADW17851	
ID	ADW17851 standard; protein; 418 AA.
XX	
XX	
AC	ADW17851;
XX	
XX	
DT	24-MAR-2005 (first entry)
XX	
XX	
DE	Pinus radiata transcription factor protein TUBBY family Seq 1629.
XX	
KW	plant; transcription; gene regulation; gene expression; transgenic plant;
KW	drought resistance; disease resistance; salt tolerance; cold tolerance;
KW	freezing tolerance; flowering; flavor enhancer; flower color.
XX	
XX	
OS	Pinus radiata.
XX	
PN	WQ2005001050-A2.

New polynucleotides isolated from plants encoding transcription factors, and polypeptides encoded by such polynucleotides, useful for regulating gene transcription and gene expression.

Claim 31; SEQ ID NO 1629; 1265pp; English.

This invention relates to novel isolated plant nucleic acid molecules, or variants thereof, that encode transcription factors. Specifically, it refers to transcription factor proteins that are capable of binding to DNA in order to regulate gene transcription and gene expression in a plants, in particular Eucalyptus grandis and Pinus radiata. The present invention describes DNA constructs containing DNA encoding a transcription factor that regulates the promoter, which is operably linked to the desired nucleic acid to be expressed. It further provides transgenic plants expressing a transcription factor that confers a trait to the plant such as increased drought, salt or disease tolerance, height change, enhanced cold/frost tolerance, enhanced color, health and nutritional characteristics, as well as improved taste, starch composition, flower longevity and germination, amongst others. Accordingly, such plants that are successfully transfected with a DNA construct can be characterized by a difference in flower color, petal or leaf shape and size, aroma or plant height. This polypeptide is a plant transcription factor protein sequence of the invention.

Sequence 418 AA;
Sequence 419 AA;
Sequence 420 AA;

Query Match	55.1%	Score 1082;	DB 9;	Length 418;
Best Local Similarity	53.2%;	Pred. No. 6.3e-112;		
Matches 226;	Conservative 59;	Mismatches 88;	Indels 52;	Gaps 8
QY	1	MTFRSLLOEM-----RSRPHRVV-----HAAASTANSDDPPS-----W	33	
DB	1	MTVKNIFODMKSGISGRSRSEFVKFURSRQSAVDDLHYRSNANKNNNSNNGGLQCSLW	60	
QY	34	SELPEELIREILIRVETVGDGWPSPRRNVVACGCRSMRILTKETIVAVPESSKLTFFI	93	
DB	61	ANMPELIRDIVIQRIESE-SSWPFKRSVACGVCMTWFQITKELVPIPELSGKLTFFI	119	
QY	94	SLKQSGPRDSLVQCFIKRNRNTQSYHYLGLTTSITDNGKFLAASKLKRATCTDIISL	153	
DB	120	SLKQPGSREPSVQCFIKRDRATSTYQLFLCLNTTLENGKFLAARKFRATSTEYIISL	179	
QY	154	RSDDISKESNAYLGRMSRNFILGKTFVFDG---SQTGAAMQKSRSSNFIKVSPRVPOG	209	
DB	180	NADDMRSRGNTYVGKLRNFIKTKPTLYIDSQPHSGAIASTRAGARRVGAKQVSPRPVAG	239	
QY	210	SYPIAHISYELNVLGSRGPRRMRCIMDTIIMSVIESRG-----VVASTSISSFSRSSPV	264	
DB	240	SYNIAHIGYELNVLGTRGPRMQCIMHSIPASAVQPGSAPTPEVPSVASLDQSLASCPT	299	
QY	265	FRSHSKPLRSNSACSGSGNNLGDPPILVNLKAPRWHEQLRCWCLNPHGRVTVASVKNFQ	324	
DB	300	FGSKCILLESSSISEPLTSGNLKDXDEPLVLNKNAPRWHEQLQCWCLNPKGRVTVASVKNFQ	359	
QY	325	LVAV-----SDCEAGQTSERILILQFGKVGKDMFTMDYGPISAFQAFACLSSPET	375	
DB	360	LVAAPEPNPSVSQNDK-----DKVILQFGKVGKDMFTMDYRPLSAFQAFACLSSFDT	413	
QY	376	RIACE 380		
DB	414	KLACE 418		
RESULT 14				
ADW17850				
ID	ADW17850	standard; protein; 418 AA.		
XX	AC			
XX	ADW17850;			
XX	AC			
XX	XX			
XX	24-MAR-2005	(first entry)		
XX	DE	birus radiata transcription factor protein TURBY family, Sec1628.		

RESULT 14	
ADWL17850	
ID	ADWL17850 standard; protein; 418 AA.
XX	
XX	ADWL17850;
XX	
XX	24-MAR-2005 (first entry)
XX	
XX	Pinus radiata transcription factor protein TUBBY family Seq 1628.
XX	
XX	plant; transcription; gene regulation; gene expression; transgenic plant;
KW	drought resistance; disease resistance; salt tolerance; cold tolerance;
KW	freezing tolerance; flowering; flavor enhancer; flower color.
XX	
XX	Pinus radiata.
XX	
XX	WO2005001050-A2.
PN	
XX	
PD	06-JAN-2005.
XX	
XX	07-JUN-2004; 2004WO-US017965.
PF	
XX	
XX	06-JUN-2003; 2003US-0476189P.
PR	
XX	
XX	(ARBO-) ARBORGEN LLC.
PA	
XX	
XX	Bloksberg LN, Bryant C, Connett MB, Emerson SJ, Frost MJ;
PI	Forster RLS, Grigor M, Higgins C, Lasham A, Lund ST, Magnus A;
PI	Phillips J, Puthigae S, Veerakone S, Westwood C, Gause K, Wood M;
XX	
XX	WPI; 2005-075542/08.
DR	
DR	N-PSDB; ADWL17064.
XX	
XX	New polynucleotides isolated from plants encoding transcription factors,
PT	and polypeptides encoded by such polynucleotides, useful for regulating

PT	gene transcription and gene expression.	PN	WO2005001050-A2.
PS	Claim 31; SEQ ID NO 1628; 1265pp; English.	PD	06-JAN-2005.
XX		XX	07-JUN-2004; 2004WO-US017965.
CC	This invention relates to novel isolated plant nucleic acid molecules, or	PF	06-JUN-2003; 2003US-0476189P.
CC	variants thereof, that encode transcription factors. Specifically, it	XX	(ARBO-) ARBORGEN LLC.
CC	refers to transcription factor proteins that are capable of binding to	XX	Bloksberg LN, Bryant C, Connett MB, Emerson SJ, Frost MJ;
CC	DNA in order to regulate gene transcription and gene expression in a	PI	Forster RLS, Grigor M, Higgins C, Lasham A, Lund ST, Magusin A;
CC	plants, in particular Eucalyptus grandis and Pinus radiata. The present	PI	Phillips J, Puthigae S, Veerakone S, Westwood C, Gause K, Wood M;
CC	invention describes DNA constructs containing DNA encoding a	XX	WPI; 2005-075542/08.
CC	transcription factor that regulates the promoter, which is operably	DR	N-PSDB; ADW18185.
CC	linked to the desired nucleic acid to be expressed. It further provides	XX	New polynucleotides isolated from plants encoding transcription factors,
CC	transgenic plants expressing a transcription factor that confers a trait	PT	and polypeptides encoded by such polynucleotides, useful for regulating
CC	to the plant such as increased drought, salt or disease tolerance, height	PT	gene transcription and gene expression.
CC	change, enhanced cold/ frost tolerance, enhanced color, health and	XX	Claim 31; SEQ ID NO 2296; 1265pp; English.
CC	nutritional characteristics, as well as improved taste, starch	CC	This invention relates to novel isolated plant nucleic acid molecules, or
CC	composition, flower longevity and germination, amongst others.	CC	variants thereof, that encode transcription factors. Specifically, it
CC	Accordingly, such plants that are successfully transfected with a DNA	CC	refers to transcription factor proteins that are capable of binding to
CC	construct can be characterized by a difference in flower color, petal or	CC	DNA in order to regulate gene transcription and gene expression in a
CC	leaf shape and size, aroma or plant height. This polypeptide is a plant	CC	plants, in particular Eucalyptus grandis and Pinus radiata. The present
CC	transcription factor protein sequence of the invention.	CC	invention describes DNA constructs containing DNA encoding a
XX		CC	transcription factor that regulates the promoter, which is operably
SQ	Sequence 418 AA;	CC	linked to the desired nucleic acid to be expressed. It further provides
	Query Match 53.5%; Score 1049.5; DB 9; Length 418;	CC	transgenic plants expressing a transcription factor that confers a trait
	Best Local Similarity 54.8%; Pred. No. 2.9e-108; Indels 31; Gaps 9;	CC	to the plant such as increased drought, salt or disease tolerance, height
	Matches 216; Conservative 57; Mismatches 90; Indels 31; Gaps 9;	CC	change, enhanced cold/ frost tolerance, enhanced color, health and
QY	11 RSRPHRVHAAASTANSSD-----PFSWSELPEELLREILIRVETDGGWPSRR 60	CC	nutritional characteristics, as well as improved taste, starch
Db	32 RRSHS--HGVAVTSGDENHCKTAADVLSHCWANLPYELLARDIIRIESSE-IWPAAR 88	CC	composition, flower longevity and germination, amongst others.
QY	61 NVVACAGVCRWRILTKEIVAVPEFSSKUTFPISLKQSGPRDSLVOCFIKRNRNTQSYHL 120	CC	Accordingly, such plants that are successfully transfected with a DNA
Db	89 NVVACAGVCRWRILTKEIVAVPEFSSKUTFPISLKQSGPRDSLVOCFIKRNRNTQSYHL 148	CC	construct can be characterized by a difference in flower color, petal or
QY	121 YGLTSLTDNCKELLAASKLRATCTDYIISLRSDISKRSNAYLGRMSNPLGTGKFTV 180	CC	leaf shape and size, aroma or plant height. This polypeptide is a plant
Db	149 YGLTSLTDNCKELLAASKLRATCTDYIISLRSDISKRSNAYLGRMSNPLGTGKFTV 208	CC	transcription factor protein sequence of the invention.
QY	181 FPGSOT---GAAKWQKRSNNFKVSPRVPQSGYPIAHISYELNLVGRGPRMRCIMDT 237	XX	
Db	209 YDSQSHSGVATSHAGRVGSKQVSPRPASVYNVAHAYELNLVGRGPRMRCIMDT 268	SQ	Sequence 414 AA;
QY	238 IPMSIVE-----SRGVASTSISFSS--RSQPVFRSHSKPLRSNASCSDSGNNLGDPP 290		Query Match 52.4%; Score 1028; DB 9; Length 414;
Db	269 IPASVVEGGLNTSTPEIPSTLVEPLASFRSS---RSKSVISESNAFSGPLNSGKLDKP 325		Best Local Similarity 53.0%; Pred. No. 7.4e-106;
QY	291 LVLSKAPRWHLQRCWCLNFGHRTVTVASVQNFQVAVSDCEAGQTSF-----RIILOFGK 346		Matches 220; Conservative 57; Mismatches 102; Indels 36; Gaps 10;
Db	326 LILKNKAPRWHLQRCWCLNFGHRTVTVASVQNFQVAVSDCEAGQTSF-----RIILOFGK 384	QY	1 MTFERSILOEM-----RSRPHRVV--HAAASTANGSDPPS-----WSEL 36
QY	347 VCKDMFTMDYGPISAFQAFALCLSSFETRIACE 380	Db	1 MSLRSIMREIKVRNGIGNISRRSFDMKVSNNHRRASQATDDTLRSVLFVQSQSWANL 60
Db	385 VCKDMFTMDYGPISAFQAFALCLSSFETRIACE 418	QY	37 PPELLREILIRVETDGGWPSRRNVVACAGVCRWRILTKEIVAVPEFSSKLTFFPISLK 96
		Db	61 PPELLHNVIQRLSEASE-GTWPARKNVVACAGVCKSWREITKEVKTPEQCKLTFFPISLK 119
		QY	97 QSGPRDSLVOCFIKRNRNTQSYHLIYGLTTS--LTNGKFLLAASKLRATCTDYIISLS 155
		Db	120 QPGRDPPMQCFIKRDRSASTYKLFGLTTPALLVENGKFLPAKKVRRHATDYIISLS 179
		QY	156 DDISKRSNAYLGRMSNPLGTGKFTVFDGS--GTGAAMQKRSNNFI--KVSPRVPQGSY 211
		Db	180 EDFSRASNTYVQKRSNPLGTFTIYDSQPPHSGAVLTSRFSRRRYSQVQSPRVPAGSY 239
		QY	212 PIAHISYELNLVGRGPRMRCIMDTIPMSIVESRGVAVASTSISGFSRSSSPVFRSHSKP 271
		Db	240 NIANIVYELNLVGRGPRMRCIMDTIPMSIVESRGVAVASTSISGFSRSSSPVFRSHSKP 299
		QY	272 LRS-NSASCSDSGNNL--GDPLVLVSNKAPRWHLQRCWCLNFGHRTVTVASVQNFQVAV 327
		Db	300 LIDFSSSSLSLEIPINVESKDVPLVKNKAPRWHLQRCWCLNFGHRTVTVASVQNFQVAV 359
		QY	328 --VSDCEAGQTSERILQFGKDMFTMDYGPISAFQAFALCLSSFETRIACE 380

Db 360 VEPSQLVTQADQKVILQFGKIGKDIFTMDYRYPPLSAFQAPALCLSSFDTKLACE 414

Search completed: December 23, 2005, 23:23:52
Job time : 202 secs